BC

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A01H 5/00, 7/00, 9/00, 11/00, C07K 7/08, 14/415, C12N 15/29, 15/63

A1

(11) International Publication Number: WO 98/37755

(43) International Publication Date: 3 September 1998 (03.09.98)

(21) International Application Number: PCT/US98/04077

(22) International Filing Date: 27 February 1998 (27.02.98)

(30) Priority Data:
60/039,063 28 February 1997 (28.02.97) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 60/039,063 (CIP)
Filed on 28 February 1997 (28.02.97)

(71) Applicant (*for all designated States except US*): THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL [US/US]; 308 Bynum Hall, Campus Box 4105, Chapel Hill, NC 27599-4105 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): DANGL, Jeffery, L. [US/US]; 208 Lake Court, Chapel Hill, NC 27516 (US). DIETRICH, Robert, A. [US/US]; Apartment BB7, 605 Jones Ferry Road, Carrboro, NC 27510 (US). RICHBERG, Michael, H. [US/US]; A6 University Gardens, 800 Pritchard Avenue Ext., Chapel Hill, NC 27516-1717 (US). EPPLE, Petra, M. [DE/US]; Apartment 167, 881 Airport Road, Chapel Hill, NC 27514 (US).

(74) Agent: BARBER, Lynn, E.; P.O. Box 6450, Raleigh, NC 27628 (US).

(81) Designated States: AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GE, GW, HU, ID, IL, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PLANT PATHOGEN RESPONSE GENE

(57) Abstract

DNA molecules encoding a family of zinc-finger DNA binding domains, which appears to function to monitor levels of a superoxide-dependent signal and negatively regulates a plant cell death pathway, including wild-type LSD1, LOL1 and LOL2, and proteins which physically interact with LSD1, indicating a function with LSD1 of controlling plant cells' response to pathogens.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

PLANT PATHOGEN RESPONSE GENE

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims benefit of U.S. Provisional Application No. 60/039,063 filed February 28, 1997.

BACKGROUND OF THE INVENTION

Field of the Invention

This invention relates to a novel DNA molecule that encodes a novel polypeptide, LSD1, which has an effect in regulating the initial response of plants to pathogens and the subsequent spread of plant cell death engendered by infection, the protein encoded by the gene, and transgenic plants comprising the DNA molecule. This invention also relates to novel DNA molecules encoding LSD1 related proteins LOL1 and LOL2. In addition, it relates to novel DNA molecules encoding proteins which directly interact with LSD1.

Description of the Related Art

Controlled induction of cell death occurs during both normal plant development and as the rapid, localized response to pathogen infection known as the hypersensitive response (HR) (Stakman, 1915; Goodman and Novacky, 1994; Dangl et al., 1996). The HR is a feature of most, but not all, disease resistance reactions in plants. The disclosure of these publications and all others cited herein, as well as of the priority application, is incorporated herein by reference.

Genetic control of disease resistance reactions is of two broad classes. The first is determined by specific interactions between particular alleles of pathogen *avr* (avirulence) gene loci and an allele of the corresponding plant disease resistance (*R*) locus. When these alleles are present in both host and pathogen, the result is disease resistance in the plant, and the interaction is said to be "incompatible". If either the plant *R* allele or the cognate pathogen *avr* gene are absent or inactive, disease results and the interaction in said to be "compatible" (reviewed by Flor, 1971; Crute, 1985; Keen, 1990; Pryor and Ellis, 1993). A great deal of progress has been made recently in understanding the molecular structure of *R* genes and their predicted products (reviewed by Dangl, 1995; Staskawicz et al., 1995; Bent, 1996). These molecules function to recognize *avr* dependent signals and trigger the plant cell to begin the chain of signal transduction events culminating in a halt of pathogen growth. The simplest mechanistic interpretation of allele-specific disease resistance is that the *R* gene product recognizes the *avr* gene product directly. Although no direct avr-R

protein interaction has been shown in planta, expression of *avr* genes in plant cells can be sufficient to trigger the HR in a *R*-dependent manner, and avr-R protein-protein interactions can occur in yeast two-hybrid systems (Gopalan et al., 1996; Scofield et al., 1996; Tang et al., 1996).

The second mode of genetic control of disease resistance is termed "non-host" resistance and describes in essence those interactions which lack genetic variability in either host or pathogen such that no virulent pathogen and no susceptible host line have been identified. While it is not beyond reason to assume that traditional "non-host" interactions are simply a series of allele specific recognition events occurring simultaneously (Whalen et al., 1988; Kobayashi et al., 1989; Valent et al., 1990), it is also possible that this mode of resistance is mechanistically distinct from that mediated by allele-specific interactions. Pathogen ligands (termed elicitors) which mediate several key non-host interactions have been isolated, although their corresponding plant receptors have not (Cosio, et al. 1992; Nürnberger et al., 1994).

Subsequent to pathogen recognition by either of these two systems, the plant cell deploys a battery of inducible defense responses. Chief among the earliest events are calcium influx, K^+-H^+ exchange leading to alkalinization of the extracellular space, and an oxidative burst (reviewed in Godiard et al., 1994; Hammond-Kosack and Jones, 1996). The latter is potentially mediated by a plasma membrane NADPH oxidase analogous to that used by mammalian neutrophils (Low and Merida, 1996), although other models exist (Bolwell et al., 1995). Parts of this cascade are linearly regulated in at least some systems: blocking of Ca^{2+} influx blocks anion channel activity, the oxidative burst and downstream events including cell death; blocking anion channels effects only ROI production and defense gene activation, but not Ca^{2+} influx (Nürnberger et al., 1994; Levine et al., 1996; May et al., 1996).

Consequent production of reactive oxygen intermediates (ROI) occurs with kinetics and magnitude suggesting a key role in either pathogen elimination, subsequent signaling of downstream effector functions, or both (reviewed by Baker and Orlandi, 1995; Low and Merida, 1996). H2O2 can have a key role in resistance responses, and cell wall strengthening (Brisson et al., 1994; Levine et al., 1994; Levine et al., 1996), and superoxide produced as the proximal ROI in the burst has also been implicated in initiating HR (Doke, 1983; Jabs et al., 1996). Transcription and translation of plant genes are required for HR. These signals are thought to culminate in transcriptional activation of a variety of plant genes, HR, and the production of both local and systemic signals that protect the plant from further infection. It is unclear whether these effector functions are controlled by linear, interdigitating, or bifurcating signal pathways.

Cell death during the HR may be a direct consequence of ROI toxicity, or it may be

a secondary consequence of signals derived from ROI. It is not known whether HR is required to halt pathogen growth. Nonetheless, HR is correlated with the onset of systemic acquired resistance (SAR) to secondary infection in distal tissue (reviewed by Ryals et al., 1996). In at least tobacco and Arabidopsis, enzymatic blocking of salicylic acid (SA) accumulation subsequent to infection alters disease resistance responses, and SA in distal tissues is required for SAR (Gaffney et al., 1993; Delaney et al., 1994; Vernooij et al., 1994). SA accumulates following the oxidative burst to high levels locally at infection sites. The biochemical properties of SA as an inhibitor of a variety of enzymes suggest a model whereby SA or a radical derived from it poisons the infected cell, causing its death (Enyedi et al., 1992; Malamy et al., 1992; Chen et al., 1994; Durner and Klessig, 1995; Rueffler et al., 1995). Recent descriptions of the morphology of cell death during infection suggest, in at least some cases, parallels with animal apoptosis (Mittler et al., 1995; Kosslak et al., 1996; Levine et al., 1996; Ryerson and Heath, 1996; Wang et al., 1996a; reviewed by Dangl et al., 1996). A molecular understanding of both the signaling events that control the onset of this specialized plant cell death and the mechanisms by which these cells die will hasten approaches to manipulate cell death to protect plants from disease.

A number of researchers have isolated mutants in Arabidopsis which exhibit constitutive onset of HR-like cell death in the absence of pathogen (Greenberg and Ausubel, 1993; Dietrich et al., 1994; Greenberg et al., 1994). These mutants resemble a variety of mutants in crop species isolated since the 1920s and broadly categorized as "lesion mimic mutations" (Langford, 1948; Kiyosawa, 1970; Walbot et al., 1983; Johal et al., 1994). A series of non-allelic mutations was isolated which expressed histochemical and molecular markers associated with disease resistance responses. These mutants subdivide the lesion mimic class into a "lesions simulating disease resistance" or *lsd* phenotype (Dietrich et al., 1994). These mutants also exhibited heightened resistance to otherwise virulent bacterial and oomycete pathogens when lesions were present, demonstrating that these cell death phenotypes can trigger pathogen non-specific resistance resembling SAR. Similar "accelerated cell death" or *acd* mutants have been described by Greenberg and Ausubel (Greenberg et al., 1994). Greenberg and Ausubel (1993) additionally isolated a mutant which though expressing an *acd* phenotype was in fact more susceptible to pathogen. It is thus possible to identify genetically at least two types of cell death, namely those which feed into a pathway culminating in establishment of a disease resistant state, and those which do not.

The *lsd1* mutant is exceptional. In conditions permissive for wild type plant growth and in the absence of detectable microscopic lesions, the *lsd1* mutant is hyper-responsive to challenge by a variety of stimuli including pathogens and low doses of chemicals which trigger the onset of SAR (Dietrich et al., 1994). Mutant *lsd1* plants are resistant to

otherwise virulent pathogens in conditions where no spontaneous cell death lesions form. Following initiation of cell death in a local spot on a leaf, lesions propagate throughout the leaf and kill it 2-4 days later. Propagation of locally initiated cell death is confined to the inoculated leaf. Thus, *LSD1* functions to negatively regulate both the initial response to pathogens and the subsequent spread of cell death. Superoxide is a necessary and sufficient trigger for this phenotype, and superoxide production precedes onset of cell death by 8-16 hours following initiation by three different triggers (Jabs et al., 1996). Therefore, the *LSD1* gene responds to either superoxide or to a signal derived from it to down regulate or dampen the cell death response, resulting in the typical locally bounded HR. The invention herein includes the *LSD1* gene, which encodes the first member of a new subclass of zinc-finger proteins in Arabidopsis.

It is therefore an object of the invention to provide a novel DNA molecule, LSD1, isolated from Arabidopsis which works to protect plant cells in response to pathogens, and DNA molecules encoding LSD1 related proteins LOL1 and LOL2.

It is a further object of the invention to provide the protein encoded by *LSD1*, and transgenic plants comprising LSD1. Knowledge of the structure of the *LSD1* gene allows accurate creation of particular mutants (e.g., deletion and point mutations), for example, mutants having a dominant negative phenotype, analogous to the mutants of Drosophila *PANNIER* gene (Ramain et al., 1993), using methods known in the art. This in turn allows engineering of transgenic crop plants which do not suffer cell death, but are still resistant to infection. In addition, expression of the dominant negative LSD1 protein may be refined so that it is expressed very quickly after infection.

The LSD1 protein is also a useful target for herbicide development. Transgenic plants may be made in which *LSD1* mutant genes are expressed which are resistant to herbicidal compounds which normally result in cell death in combination with the wild-type LSD1. Mutants of the *LSD1* gene are tested in a *lsd1* background to determine if the mutant has a normal or novel function, and in a wild-type background to determine the existence of a dominant negative function.

Other objects and advantages will be more fully apparent from the following disclosure and appended claims.

SUMMARY OF THE INVENTION

The invention herein comprises the DNA molecule of the wild-type LSD1, which functions to monitor levels of a superoxide-dependent signal and negatively regulates a plant cell death pathway. The predicted LSD1 protein contains three zinc-finger domains, defined by CxxCxRxxLMYxxGASxVxCxxC (SEQ ID NO:54). The invention further comprises a protein encoded by LSD1, and transgenic plants comprising LSD1, and

mutations thereof.

In particular, the preferred embodiments of the invention herein include the following: an isolated DNA molecule, encoding the LSD1 polypeptide sequence, selected from the group consisting of SEQ ID NOS:13-15; the LSD1 DNA molecule having the nucleotide sequence as set forth in SEQ ID NO:13; the DNA molecule that is cDNA; the DNA molecule which is genomic DNA; a chimeric construction comprising a promoter sequence and the LSD1 DNA molecule or portions of the LSD1 DNA molecule; a recombinant plant transformed with the LSD1 DNA molecule; a transformed plant comprising a DNA molecule encoding a protein as set out in SEQ ID NO:16 or SEQ ID NO:17; an isolated protein molecule comprising the protein set out in SEQ ID NO:16 or SEQ ID NO:17; a transformation vector comprising a LSD1 DNA molecule as set forth herein; an isolated DNA molecule encoding the zinc finger consensus sequence shown in SEQ ID NOS: 1-3; and anything that hybridizes to the LSD1 DNA molecule set forth herein under hybridization conditions as defined herein.

Other objects and features of the inventions will be more fully apparent from the following disclosure and appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-C show the physical delineation of the *lsd1* mutation. Figure 1A shows YAC clones at *lsd1*. The arrowheads imply the YAC clone extending in the direction given, solid vertical black bars denote YAC ends used to isolate genomic phage clones and subsequently converted into CAPS RFLP markers as described (refer to Figure 2 for their map position and to Tables 1 and 2, Examples II and III, for their definition). Figure 1B shows the three BAC clones which contained the CAPS markers listed above BAC1G5. The arrowheads imply extension of the BAC clone in the direction shown. The scale in Figures 1A and 1B are the same. Figure 1C shows the genomic phage clones positioned under an expansion of three of the BACs. The diamond-filled bar represents the 8A6-1.3 clone, which co-segregated with *lsd1*, used to isolate these phage. The *lsd1* deletion is noted at the bottom.

Figure 2 is a genetic linkage map of the *lsd1* region. The vertical line at the left represents the section of Arabidopsis chromosome 4 between CH42 and B9-1.8 (telomeric toward bottom). CAPS-based RFLP markers discussed in the text intersect the chromosome, and their relative recombination frequencies in the F_2 mapping population are placed in the center. The number of meioses identified among the total number of F_2's scored is at the right. The arrowhead denotes the co-segregating marker.

Figures 3A-C show molecular fine mapping of the *lsd1* locus. Figures 3A and 3B show genomic DNA blots demonstrating the presence of a 0.8 kb deletion om the *lsd1*

mutant. Genomic DNA (5 g) from wild type Ws-0 or *lsd1* was digested with (for each pair of lanes from left to right) EcoRI, HindIII, a double digest of HindIII and Xbal, or Kpnl. In Figure 3A, the blot was probed with the 0.8 kb EcoRI-Xbal. In Figure 3B, a duplicate blot was probed with the 4.5 kb Pstl-Xhol fragment. The probes are depicted in Figure 3C, and were isolated from phage clones depicted in Figure 1C. Molecular weight markers are the Gibco-BRl 1 kb ladder. Figure 3C shows the restriction map in and around the *lsd1*gene. The extent of the deletion of this locus is shown as are the extent of the hybridization of the various restriction fragments with *lsd1* cDNAs. Genomic restriction fragments used in complementation experiments are underlined. The asterisk refers to an Xhol site derived from the phage lambda cloning junction.

Figure 4 shows that the *lsd1* mutation is an mRNA null allele. RNA blots (1 g of polyA+ RNA) from leaf tissue of 5 week old plants kept in short days (permissive for *lsd1* growth) 3 days after spraying with either INA (0.3 mg/ml powder containing 25% active ingredient, or 4 mM), or wettable powder control. Spreading *lsd1* lesions had just started to appear at the time of leaf harvest. Probes were purified inserts from the LSD1 cDNA as represented by EST 82D11T7 (top), a PR-1 cDNA (Uknes et al, 1993b), and an actin cDNA. The blot was probed successively in the order displayed.

Figure 5 shows the zinc finger domains (SEQ ID NOS:1-3) of the predicted LSD1 protein and the alignment of the three zinc finger domains. The numbers at the left and right refer to amino acid residue position in the deduced LSD1 protein. Vertical lines indicate conservation in pairwise comparison, and a colon indicates conservative substitution. A consensus sequence is listed below, with conservative substitutions noted in the second line of consensus where "+" is basic, plus charged; and "@" is amide, polar, uncharged, hydrophilic.

Figure 6 shows how the carboxyl portion of the deduced LSD1 protein is related to known DNA-binding and transcription factors. Vertical lines indicate conservation in pairwise comparison, and a colon indicates conservative substitution. Figure 6A shows homology of a slightly longer portion of the deduced LSD1 protein with mammalian insulin receptor substrate proteins. The LSD1 translation product (SEQ ID NO:4) is shown on the top, aligned with the mouse insulin receptor substrate (SEQ ID NO:5). In this region, all mammalian insulin receptor substrates are identical. Figure 6B shows the homology of LSD1, on each top line, with four known transcription factors. The LSD1 translation product (SEQ ID NO:6) is shown on top, and below it are the related domains from a human early growth response (EGR) Zn-finger protein (SEQ ID NO:7, a human TGF--early induced Zn-finger protein (SEQ ID NO:8), a *Xenopus laevis* H-L-H transcription factor (SEQ ID NO:9), and the human ELK-1 protein (SEQ ID NO:10). Figure 6C shows the homology of a LSD1 transcription product (SEQ ID NO:11) with a putative maize

transcription initiator binding protein (SEQ ID NO:12). GenBank accession numbers of each protein are listed at the right.

Figure 7 shows the consensus sequence of the zinc finger domains (SEQ ID NOS:63-65, respectively) of LSD1 (A), LOL1 (B) and LOL2 (C).

Figure 8 shows the homologies between the first (A), second (B) and third (C) zinc finger domains of LSD1, LOL1 and LOL2

DETAILED DESCRIPTION OF THE INVENTION AND PREFERRED EMBODIMENTS THEREOF

The present invention provides a genomic DNA sequence (SEQ ID NO:13) and a cDNA sequence (SEQ ID NOS:14-15) or the *LSD1* gene which is required for the regulation of initial plant response to pathogens, and cDNA proteins deduced (from short form, MG7-SEQ ID NO:16; from long form, MG, SEQ ID NO:17).

In addition, the invention herein provides functional protein domain sequences involved in regulating genes controlling cell death. Gene expression can be regulated by attaching a promoter to the LSD1 gene, which may be either the native promoter or any other promoter.

The invention herein includes the DNA molecule having the nucleotide sequence as set forth in SEQ ID NOS:13, 14 and 15, encoding either of two *LSD1* polypeptides, which are preferably the *LSD1* polypeptides set forth in SEQ ID NOS:16 and 17. This DNA molecule may be cDNA or genomic. The invention also includes as the open reading frame any chimeric construction comprising a promoter sequence and the DNA molecule of the invention, a recombinant plant transformed with the DNA molecule, and any transformation vector comprising the DNA of the invention. In addition, the DNA sequence of either the full-length SEQ ID NO:13, or a shortened or otherwise modified version thereof, may be modified to optimize its expression in plants, with codons chosen for production of the same or a similar protein as encoded by the wild type *LSD1* gene. Other modifications of the *LSD1* gene that yield a protein having essentially the same properties as the *LSD1* gene are included within the invention herein.

The invention herein also includes anything that hybridizes to the LSD1 DNA (SEQ ID NO:13) of the invention as discussed above, under hybridization conditions, which are defined as: 7% Na dodecyl sulfate (SDS), 0.5 M sodium phosphate, pH 7.0, 1 mM EDTA at 50C, and wash in 2X SSC buffer, 1% SDS, at 50C (Church and Gilbert, 1984). Proc. Natl. Acad. Sci. USA 81:1991-1995 (1984)).

The novel *LSD1* gene of the present invention, it its wild type form or as mutated by selected mutations and genetically engineered derivatives obtained as is known in the art,

and proteins encoded thereby, are included in the invention herein, and may be transferred into any plant host using methodology known in the art for purposes of altering the extent and type of plant resistance to pathogens, and to change resistance to particular herbicides.

The mutant phenotype of the null *lsd1* allele suggests that the wild type product is a negative regulator of cell death. In addition, *lsd1* reacts to both nominally virulent pathogens, and to chemicals which trigger the onset SAR, with an HR-like response. But it is important to note that *lsd1* expresses wild type timing of *R* gene driven HR (Dietrich et al., 1994)--it is the subsequent spread of cell death which distinguishes the mutant. Thus, cell autonomous signals required for *R* gene function are intact in an *lsd1* null, but the response to cell non-autonomous signals emanating from cells undergoing HR is perturbed. Collectively, these features of the mutant phenotype suggest that LSD1 functions to limit both the initiation of defense responses and the subsequent extent of the HR. The fact that an *lsd1* null is hyper-responsive to signals initiating the defense response and HR-like cell death additionally suggests that these pathways are functionally intact in the wild type cell, but require a threshold level of signal for full activation.

LSD1 appears to act as a transcription factor (or as a protein which sequesters a transcription factor). As outlined above, the oxidative burst in an infected cell generates a superoxide-dependent signal up-regulating the HR pathway. This signal overcomes the negative regulatory function of the available LSD1, and drives primary responding cells into the HR pathway. Additionally, the cells undergoing HR amplify the signal, probably via a sustained extracellular oxidative burst, to neighboring cells. The primary signal molecule may be diffusible over short ranges (Levine et al., 1994), could act as an autocrine signal, and could lead to the accumulation of a secondary signal molecule in a steep spatial gradient from the infection site. At a critical point in the signal gradient, a threshold is reached. Above that point the pro-death pathway operates, and below it the pro-death response would be attenuated by LSD1. Such a gradient is formed by SA and SA-conjugates (Enyedi et al., 1992); SA biosynthesis can be induced by hydrogen peroxide (Leon et al., 1995); and sub-effective doses of SA can amplify pathogen-derived signals (Kauss et al., 1992; Kauss and Jeblick, 1995; Mauch-Mani and Slusarenko, 1996). Thus, it could be that an SA gradient dictates LSD1 activity.

Constitutive expression levels by LSD1 could suffice to protect cells below the critical signal threshold for death induction. The time lag of 12-16 hours observed between superoxide production initiated in *lsd1* by a variety of triggers and the onset of cell death (Jabs et al., 1996), which could provide sufficient time for up-regulation of LSD1 activity before irrevocable commitment to death during wild type responses, so that cell death could spread until sufficient active LSD1 accumulates. Alternatively, this time lag could represent a requirement for biosynthesis of pro-death intermediates and LSD1 normally could operate

by interdicting this pathway. LSD1 could positively regulate anti-cell death targets, potentially including genes involved in cell survival, ROI de-toxification, or in degradation of a key intermediate in the pro-death pathway. Alternatively, LSD1 could act as a transcriptional repressor directly on genes in the pro-death effector pathway. This scenario differs from the first only in that the set of target genes would be different. The availability of extragenic suppressors of *lsd1* will aid in identifying LSD1 targets (Jabs et al., 1996).

This model also explains the runaway cell death phenotype of the null *lsd1* mutant. In the absence of LSD1, the threshold normally required before commitment to HR is removed. Thus, minimal up-regulation of the superoxide-dependent signal drives the cell into the HR pathway. Hence the ability of *lsd1* to respond to virulent pathogens as if resistant derives from a lack of background inhibition of the HR pathway normally operating in the cell. Moreover, extracellular superoxide produced during the oxidative burst initiates the same series of events in cells immediately surrounding the site of initiation, and the cell death propagation indicative of the *lsd1* phenotype results. Because the null *lsd1* mutant still requires superoxide for initiation of cell death propagation, it is unlikely that superoxide directly regulates LSD1 activity. This further suggests that a superoxide-dependent signal is the autocrine which propagates the response to neighboring cells.

The *A. thaliana lsd1* mutant phenotype is characterized by enhanced disease resistance, spontaneous formation of lesions in the absence of cell death initiators and failure to limit the extent of cell death. The wildtype LSD1 protein therefore negatively regulates a cell death pathway involved in plant defense responses.

The *LSD1* gene encodes a protein containing a novel zinc finger protein, which is included in the invention herein and is defined by its three consensus zinc fingers: CxxCRxxLMYxxGASxRxVxCxxC (SEQ ID NO:52). These three zinc finger domains have not been observed before in the range of zinc finger proteins. As shown in Dietrich et al., 1997, the *LSD1* gene is a key negative regulator of hypersensitive cell death in plants. We sought other versions of this consensus zinc finger sequence in other plant proteins.

The data on homologies between the LSD1 and LOL1 and LOL2 zinc finger domains indicates that LSD1 as well as LOL1 and LOL2 are members of a novel subclass of zinc finger proteins that are involved in plant cell death pathways. LOL1 and LOL2 might function in cell death phenomena leading to hypersensitive response and disease resistance as has been shown for LSD1. The homologues may also be involved in programmed cell death (PCD) pathways occurring in plants. Examples of PCD n plants include lateral root development, tracheary element differentiation, and abscission of leafs. Preliminary expression studies suggest that LOL2 is expressed in flowers and siliques. Thus a role for LOL2 in PCD pathways leading to petal senescence, anther dihiscence or

PCD of nucellar cells is not unlikely. It is also possible that LOL2 is involved in the hypersensitive response and disease resistance in flowers, thus protecting seeds and ultimately the following generations from pathogen. Alternatively, LOL2 could be up-regulated during the hypersensitive response. Use of *LOL1* and *LOL2* should allow prediction of the protein's function with respect to protection from programmed cell death.

The consensus sequences defined by the LSD1, LOL1 and lOL2 zinc finger domains (Figures 7-8) are thus far unique in the available deduced protein databases. Because zinc finger domains of this type bind DNA and thereby regulate gene activation, it is highly likely that the consensus zinc finger domain defined here is required for proper regulation of related sets of genes. Furthermore, because zinc finger DNA binding domains of related sequence generally control related cellular processes, the new consensus defined here should also do so. Because LSD1 is known to negatively regulate cell death induced by pathogens, it is highly likely that LOL1 and LOL2 also control plant cell death. Thus, the utility of this portion of the invention lies in production of transgenic plants which have mutated versions of the *LOL1* or *LOL2* genes or which overexpress these proteins. Such plants will likely be more resistant to pathogen attack, if, in the first case, the *LOL* genes function to repress defense response (as does *LSD1*). Alternatively, if the *LOL* genes function to activate defense mechanisms, then overexpression will lead to a more effective pathogen response. Because zinc finger proteins featuring other non-LSD1 type DNA binding domains function to either activate or repress gene transcription, we cannot distinguish at present between these two models.

The invention also includes plant proteins, and the genes which encode them, which directly interact with LSD1 protein. Gene regulation in response to pathogen attack is controlled, in part, by the repression and activation of genes. The LSD1, LOL1 and LOL2 proteins encode a novel branch of the zinc-finger DNA binding protein superfamily with roles in controlling plant cell death. As such, they are expected to interact with other proteins. Paradigms of gene activation currently demonstrate that DNA binding proteins can have two classes of "partners". The first class sequesters the DNA binding protein in the cell's cytosol. These partner proteins hold the DNA binding protein out of the nucleus until the correct cellular stimulus is received. This stimulus disrupts the physical interaction, and the DNA binding protein is free to migrate into the nucleus and activate or repress transcription. The second class of protein which interacts with DNA binding protein is made up of proteins which are partners having the role of "enhancing" the gene activation or repression encoded by the DNA binding protein. These partners are termed "co-activators" or "co-repressors" and they may or may not have intrinsic DNA binding activity. We have identified several genes whose protein products interact physically with the LSD1 protein using a common assay, called a "yeast two-hybrid interaction trap" to detect such

interactions genetically (Fields and Sternglanz, 1994; Finley and Brent, 1996). Because the inactivation of LSD1 by mutation leads to enhanced disease resistance, the LSD1 partner proteins represent novel targets for engineering plants with enhanced resistance to pathogens. Thus, this invention includes all proteins which interact with the cell death regulator LSD1 (SEQ ID NOS: 66-91 (includes sequential pairs of nucleic acids and corresponding amino acid sequences).

The features of the present invention will be more clearly understood by reference to the following examples, which are not to be construed as limiting the invention.

EXAMPLES

Example I **Care and maintenance of plants**

Plants were grown in a chamber at 9 hours light per day, 22°C day temperature and 20°C night temperature essentially as described (Dietrich et al., 1994).

Example II **Isolation of DNA and RNA, probe preparation, cloning**

Small scale genomic DNA preps were made from single leaves (~1cm long rosette leaves) (Lukowitz et al., 1996). The DNA pellet was re-suspended in 50 ml of Tris/EDTA (TE) and 1 ml was used in a 20 ml polymerase chain reaction (PCR). Large scale genomic DNA preps were done based on the protocol of (Rogers and Bendich, 1985), modified such that concentration in the 2X hexadecyltrimethylammonium bromide (CTAB)(Sigma, St. Louis, MO) buffer was increased to 3% and the precipitated DNA was resuspended in Tris/EDTA/sodium chloride (TEN) buffer and digested with 100 mg/ml, followed by two extractions with chloroform/iso-amyl alcohol and a final precipitation.

RNA was isolated by grinding fresh tissue in liquid nitrogen to a fine powder and extraction in 1 ml of Trizol reagent (Gibco-BRL, Gaithersburg, MD) per 100 mg tissue fresh weight. RNA was isolated according to the manufacturer's protocol. PolyA+ RNA was isolated using DynaBeads (Dynal, Oslo, Norway). RNA blots were formaldehyde agarose gels and contained either 15 mg total RNA or 1 mg polyA+ RNA. HyBond filters for DNA or RNA blots (Amersham, Little Chalfort, United Kingdom) were hybridized in 6xSSC, 5X Denhardt's solution, 0.1% SDS and 100 mg/ml sheared Herring sperm DNA at 65°C. Washes were in 0.2X SSC, 0.1% SDS at the same temperature. RNA blots were stripped for re-hybridization in 5 mM TRIS/2mM EDTA, (pH8.0), 0.1X Denhardt's solution for 1 hour at 65°C.

Example III **Isolation of new CAPS markers and genetic mapping of *lsd1***

After establishing linkage to the agamous (*AG)* co-dominant amplified polymorphic sequences (CAPS) marker (Konieczny and Ausubel, 1993), we subcloned and end-sequenced a 1.6 kb HindIII fragment from the RFLP cosmid marker g3883 (position 73.5 on the Arabidopsis RI map; Lister and Dean, 1993; see http://nasc.nott.ac.uk/RI_data/top_frame.html), and primers designed based on this sequence. This primer set amplified a rapid amplified polymorphic DNA (RAPD) marker (size difference in Ws-0 versus Col-0 without restriction digestion), and map data generated using this primer allowed us to place *lsd1* below (telomeric to) it. Probe B9-1.8, isolated as a 1.8 kb SstI-EcoRI fragment from the JGB9 genomic phage clone (RI map position ~75; gift of Dr. George Coupland, Cambridge Laboratories, Norwich U.K.) was converted into a CAPS marker. Mapping of this polymorphism placed *lsd1* above (centromeric to) it (Fig. 2). Recombinants were identified as homozygous for one of these CAPS markers, and heterozygous for the other using DNA from F2 individuals. F3 progeny from these recombinants were then scored as either homozygous *lsd1*, segregating *lsd1*, or homozygous wild-type for lesion spread. All CAPS markers we developed are described in Table 1 (below).

Table 1. New PCR based RFLP (CAPS) markers derived during cloning of *lsd1*

Marker	Enzyme	PCR prod.	Col-0	Ws-0
ch42	ClaI	1.4 kb	750 bp	1.4 kb
			650	
g3883-1.6	none		1.4 kb (uncut)	0.7 kb (uncut)
g13838-1.4	HinfI	1.4 kb	450 bp	450 bp
			330	330
			280	280
			200	160
B9-1.8	HinfI	1.8 kb	420 bp	420 bp
			260	260
			240	
			180	180
				160
			140	140
1H1L-1.6	DdeI	1.6 kb	1.0 kb	700 bp
			300 bp	300 (doublet?)
5F7R-1.5	NlaIV	1.5 kb	1.0 kb	1.2 kb

			250 bp	250 bp
			200 bp	
20B4-1.6	Ddel	1.6 kb	900 bp	700 bp
			400	400
				220
			180	180
8A6-1.3	Taql	1.3 kb	800 bp	800 bp
			400	250
			220	150

Example IV **Map refinement**

YACs were defined (Schmidt et al., 1995; Schmidt et al., 1996, http://genome-www.stanford.edu/Arabidopsis/JIC-contigs.html), confirmed by DNA blotting to establish a contig and their ends were isolated by vectorette PCR as described (Matallana et al., 1992; Grant et al., 1995). These ends were also used to isolate genomic phage from a Ws-0 genomic library (Fig. 1). Insert fragments of 1-3 kb were cloned into PBS and end sequenced for derivation of primers identifying new CAPS. PCR conditions (DNA Engine MJ Research) for all CAPS primer pairs except 8A6-1.3 and *lsd1* deletion primers are: 92°C, 3'; 35 cycles of (denature 92°C, 30"; anneal 50°C, 30"; extend 72°C, 2'30"); 72°C, 3'. For 8A6-1.3 and the *lsd1* deletion primer pairs we used 53°C annealing. Table 2 shows the primer sequences used to identify new CAPS markers.

Table 2. Primer sequences used to identify new CAPS markers used for cloning *lsd1*

ch42 for	5'-cag tgg atc ttt cct cag acg-3' (SEQ ID NO:18)
ch42 rev	5'-cat ctt ctt ctg caa tct ggg-3' (SEQ ID NO:19)
g3883-1.6 for	5'-cat cca tca aac aaa ctc c-3' (SEQ ID NO:20)
g3883-1.6 rev	5'-tgt ttc aga gta gcc aat tc-3' (SEQ ID NO:21)
g13138-1.4 for	5'-cac gtt agt tag tta gaa gg-3' (SEQ ID NO:22)
g13138-1.4 rev	5'-ctg atg ttc tct aca aat gg-3' (SEQ ID NO:23)
B9-1.8 for	5'-cgt atc cgc att tct tca ctg c-3' (SEQ ID NO:24)
B9-1.8 rev	5'-cat ctg caa cat ctt ccc cag-3' (SEQ ID NO:25)
1H1L-1.6 for	5'-ttg agt cct tct tgt ctg-3' (SEQ ID NO:26)

1H1L-1.6 rev	5'-cta gag ctt gaa agt tga tg-3' (SEQ ID NO:27)
5F7R-1.5 for	5'-gaa tgg tgt aac caa act c-3' (SEQ ID NO:28)
5F7R-1.5 rev	5'-cat acc gta tga tgg aac-3' (SEQ ID NO:29)
20B4L-1.6 for	5'-gaa ctc att gta tgg acc-3' (SEQ ID NO:30)
20B4L-1.6 rev	5'-cta aga tgg gaa tgt tgg-3' (SEQ ID NO:31)
8A6-1.3 for	5'-cca aga aga gaa aac gga ga-3' (SEQ ID NO:32)
8A6-1.3 rev	5'-aac aat agg agg tgc aga gt-3' (SEQ ID NO:33)
Primers to amplify across the *lsd1* deletion:	
lsd1 far side:	5'-acc taa caa aaa gaa aag tgt gtg agg-3' (SEQ ID NO:34)
lsd1 outside	5'-ata ata aac cct act agc tct aac aag-3' (SEQ ID NO:35)
lsd1 alt. spl. 5'	5'-ctg cta ctt tca tcc aaa c-3' (SEQ ID NO:36)

Example V **Vector construction for complementation**

The Agrobacterium vacuum infiltration procedure was used to generate transgenic plants (Bechtold et al., 1993; Grant et al., 1995). Vectors were derived from pGPTV-Hyg (Becker et al., 1992) as follows: pSGCGF was made by restricting pGPTV-Hyg with HindIII and SacI and replacing this fragment with a HindIII-SacI fragment containing the polylinker from pIC20H (GenBank accession L08912; provided by Steve Goff, Novartis, Research Triangle Park, N.C). Either the 7kb XhoI or 4.5 kb PstI-XhoI genomic fragments were cloned into this, the former into the unique vector SalI site, the latter as a SacI-SalI fragment derived from an intermediate cloning step into pBS as a PstI-XhoI fragment. The pHyg35S vector was made by cloning a four enhancer-containing 35S promoter fragment as a HindIII-XbaI fragment into pGPTV-Hyg (provided by Dr. Douglas C. Boyes, Univ. of North Carolina, Chapel Hill). The EST 82D11 cDNA sequence was isolated as a SalI-XbaI fragment from pZL1 (Newman et al., 1994) and cloned into XhoI-XbaI digested pHyg35S.

Example VI **Cloning**

The genomic Ws-0 library in lGEM11 was a gift of Dr. Kenneth A. Feldmann (Univ. of Arizona). The cDNA library is an oligo-dT primed library prepared from polyA+ Col-0 mRNA from leaves cloned into lZAPII (Stratagene, La Jolla, CA) according to the manufacturer's instructions (gift of Dr. Douglas C. Boyes and Dr. Murray R. Grant).

Example VII ***LSD1* sequences**

The sequences of the *LSD1* cDNA (SEQ ID NOS:14 and 15) and the 4.5 kb *LSD1* XhoI-PstI genomic fragment (SEQ ID NO:13; the longest 5'LSD1 cDNA starts at base 1892 of this sequence) are deposited in GenBank as accessions U 87833 and U 87834, respectively. Endpoints of the various *LSD1* cDNAs isolated are shown in Table3A and examples are provided by SEQ ID NO: 14 (short form from cDNA MG7 as shown in Table 3) and SEQ ID NO:15 (long form, from cDNA MG8). The polypeptides deduced from these are shown in Fig. 11-12, respectively. Table 3B shows the sizes of each intron deduced from comparison of the sequence shown in SEQ ID NO:13.

Table 3. Sequence characteristics of the LSD1 gene

Endpoints of independent LSD1 cDNAs

cDNA	5' end point	Alternate splice	3' end point
MG7(2)	C 1	short	A 1021
EST 82D11	A 27	short	T 1031
MG4	C59	short	A 1188*
MG10	C 59	short	G 1225
MG5	G 67	short	A 1205
MG2 (4)	G 90	short	A 1106
MG8 (2)	G 98	long	A 1082
MG16 (2)	C 103	short	A 1066
MG11	C 117	long	G 1225

Numbers in parentheses refer to the number of isolates of the same clone. Nucleotide numbers at the 5' and 3' ends refer to nucleotide positions from SEQ ID NO:13. An A at the 3' endpoint can be either an A in the genomic sequence or the first A of the polyA tail. The endpoint marked with an * had no polyA tail.

Intron sizes

intron #	size in nucleotides
1	88
2 (short splice)	68
2 (long splice)	129
3	89
4	489
5	100
6	92

7	87

Intron splice junction positions are located at bses 198-199, 260-261, 447-448, 552-553, 692-693, 764-765, and 836-837 in SEQ ID NO:13.

Example VIII **Genetic and physical mapping of *lsd1***

The *lsd1* mutation segregates as a monogenic recessive (Dietrich et al., 1994). F2 progeny of a cross between *lsd1* (Ws-0 background) and Col-0 (*LSD1*) were analyzed using the co-dominant amplified polymorphic sequences (CAPS) mapping procedure (Konieczny and Ausubel, 1993) to first establish linkage to the *AG* marker on chromosome 4. The closely linked g13838 probe (3 recombinants in 1632 meioses) was used to identify YAC (yeast artificial chromosome) clones (Schmidt et al., 1995; Schmidt et al., 1996). We constructed a physical contig of these YACs, shown in figure 1A. We used labeled YAC ends CIC1H1L, yUP5F7R and EG20B4L to isolate genomic phage clones, subcloned fragments form each of these, end-sequenced the subclones, derived primer sequences and developed new CAPS markers (see Tables 1 and 2). The CAPS markers 1H1L-1.6 and 5F7R-1.5 mapped closest to *lsd1* (1 and 3 recombinants, respectively from 2054 meioses); see Tables 1 and 2 for new CAPS markers). We hybridized these two CAPS markers to filters containing bacterial artificial chromosome (BAC) clone arrays (Choi et al., 1995, distributed by the Arabidopsis Biological Resource Center, Ohio St. Univ.), and isolated the five BAC clones depicted in Figure 2B. Because 5F7R-1.5 and 1H1L-1.6 genetically flank *lsd1* (Figure 1B), BAC clone 1G5 should contain the gene.

As 1G5 was the only BAC clone to physically span the relevant genetic region, we connected BACs 6H3 and 8A6 by walking in a genomic phage library. We defined a 5kb HindIII fragment from BAC 8A6 which hybridized only to itself and BAC 1G5. When used as a probe on filters containing restriction digests of the relevant BAC clones, this fragment hybridized to a 1.3 kb EcoRI fragment which also was present only on BACs 8A6 and 1G5. This 8A6-1.3 clone, (small box in Figure 1C) was used to isolate three phage clones, two of which are depicted in Figure 1C. Labeled inserts from each detected BAC clones 1G5, 6H3 and 8A6, thus providing multiple redundancy of genomic cloned DNA encompassing *lsd1*. We also converted 8A6-1.3 into a CAPs marker, and found that it co-segregated with *lsd1* in 2054 meioses. This map resolution of approximately 0.05 map units, suggested that *lsd1* was within 5-15 kb (at 100-300 kb per map unit; Schmidt et al., 1995; Schmidt et al., 1996) in either direction of 8A6-1.3.

We probed genomic Arabidopsis DNA blots of digested wild type Ws-0 and *lsd1* to confirm co-linearity of the cloned and genomic DNA immediately surrounding 8A6-1.3. We noted that a variety of fragments detected a genomic DNA rearrangement in *lsd1* relative to wild type Ws-0 (data not shown). This rearrangement corresponded to a loss of

restriction sites and a deletion as noted in Figures 1C and 3C. The *lsd1* mutant comes from an Agrobacterium mutagenized population of Arabidopsis, and it is known that the transformation procedure can generate non-T-DNA associated mutations (Feldmann, 1991). We subcloned and sequenced various wild type genomic DNA fragments at this position, and compared their sequences to several databases, including the Arabidopsis EST database (Rounsley et al., 1996, http://www.tigr.org/tdb/at/at.html). One EST clone (EST 82D11T7; GenBank accession T45220) exhibited blocks of identity to our genomic DNA sequence, suggesting the presence of introns in the latter. Because the gene encoding this EST is largely deleted in *lsd1*, it became a candidate *LSD1* gene.

Example IX **Complementation of *lsd1***

To confirm that the genomic deletion encompasses *LSD1*, we constructed subclones from the genomic phage as shown in Figure 3C for complementation into the T-DNA binary vector pSGCGF. Because the typical method for generation of transgenic Arabidopsis, vacuum infiltration of Agrobacterium carrying binary T-DNA vectors, triggers the propagative cell death indicative of the *lsd1* phenotype, we devised an alternate complementation strategy. We transformed F1 plants of *lsd1* x Col-0, and plated surface-sterilized seeds of the next (F2) generation onto media containing hygromycin as a selective antibiotic. We then identified hygromycin resistant transformants which were homozygous for Ws-0 alleles at 5F7R-1.5, 1H1L-1.6, and 8A6-1.3, and thus were *lsd1/lsd1* homozygous mutants. These individuals contained both mutant and wild type alleles for the CAPS marker which spans the *lsd1* deletion, because a wild type allele is present on the transgene. These transgenic plants were treated with droplets of 2,6-dichloroisonicotinic acid (INA); 0.3 mg/ml wettable powder containing 25% active ingredient, Uknes et al., 1993a) a potent inducer of SAR and the *lsd1* phenotype (Dietrich et al., 1994). If the mutation were complemented, then INA treatment should not lead to spreading cell death. Table 4 shows that transgenic plants carrying either the 7kb XhoI fragment or the 4.5 kb PstI-XhoI (Figure 3C) all survived this treatment, and are thus complemented for the *lsd1* mutation. Selfed F3 progeny from a complemented F2 individual carrying either the 4.5 kb XhoI-PstI fragment or the 7 kb XhoI fragment were also analyzed. All F3 progeny which inherited the transgene were complemented (Table 4), while all of their non-transgenic sibs still exhibited the *lsd1* phenotype (data not shown). In no case did wild type control plants exhibit spreading cell death after INA application.

Table 4. Complementation of the *lsd1* mutant

	# of plants complemented/# transgenics tested from:	
Construct	Independent F2s	Transgenic F3 progeny

7 kb XhoI	1/1[A]	20/20[B]
	3/3[C]	21/21[C]
kb PstI-XhoI	2/2[A]	14/14[B]
35S-cDNA	1/1[A]	19/19[B]

[A] Selected for hygromycin resistance and screened for homozygous Ws-0 alleles through the *lsd1* genetic interval as described, except where noted in [C]. Individual F_2s were both drop tested with INA and shifted to LD conditions (Dietrich et al., 1994).

[B] Selfed progeny from a complemented F_2 individual (homozygous Ws-0 alleles through the *lsd1* interval) were screened by PCR at F_3 for presence of the hygromycin resistance gene and then INA tested.

[C] F_2 parents were identified as hygromycin resistant and heterozygous through the *lsd1* interval, then selfed and re-screened as hygromycin resistant and homozygous Ws-0 through the *lsd1* interval at F_3 before INA testing.

Due to low numbers of independent F2 transformants which were homozygous mutant through the *lsd1* interval from the original transformation, we also isolated F2 transformants carrying the 7 kb XhoI fragment which were originally identified as heterozygote at the CAPS markers flanking *lsd1*. Selfed progeny from these should segregate both the transgene and the *lsd1* mutation. Among these progeny, we identified F3 individuals which were homozygous Ws-0 through the *lsd1* interval and carried the transgene. As shown in Table 4, these also were all complemented for protection against INA-induced spreading cell death. We conclude that the 4.5kb PstI-XhoI fragment carries the *lsd1* gene and sufficient *cis* control elements to ensure its expression.

All transgenic plants complemented for the INA-induced *lsd1* mutant phenotype were also complemented for initiation of spreading cell death after transfer to non-permissive long day conditions as well (Dietrich et al., 1994; not shown). Thus, the complementing DNA corrects the mutant phenotype induced by two independent stimuli.

<u>Example X</u> **Identification of alternately spliced *LSD1* transcripts**

We sequenced all of the complementing 4.5 kb PstI-XhoI genomic DNA fragment (SEQ ID NO:13), eight independent cDNAs (Example VII) and completed the sequence of the full 82D11T7 EST sequence. Among the cDNAs, we identified two classes expressing open reading frames of either 184 or 189 amino acids (SEQ ID NO:16 and 17). An alternate splice which adds 61bp to the 5' region of some cDNAs also provides an alternate translation start, hence, the extra five amino acids in SEQ ID NO:17. The sequences of both cDNA classes matched exactly the genomic sequence except at the positions of 7 introns

(see Table 3). Nucleotide 1 of the longest cDNA is at position 1892 in the 4.5 kb PstI-XhoI genomic sequence (SEQ ID NO:13). Thus, 1891 nucleotides of promoter are sufficient for appropriate expression in complementation of the *lsd1* mutation. The cDNA 5' ends are clustered (Table 3), suggesting that the longest could be full length. We also complemented the *lsd1* mutation by transformation of the full insert from EST clone 82D11T7 expressed from the strong and constitutive cauliflower mosaic virus 35S promoter (see Table 3) proving that this cDNA contains the entire LSD1 coding region. The 3' ends of these cDNAs are very heterogeneous, suggesting the presence of multiple polyadenylation addition signals (Table 3). No other significant open reading frames were observed in the 4.5 kb PstI-XhoI genomic clone.

When either the EST 82D11T7 clone, or a 0.8 kb EcoRI-XbaI genomic fragment covering the *lsd1* deletion were used as probes on RNA blots, a rare mRNA of approximately 1.2 kb was detected in leaf tissue of wild type Ws-0 plants (Figure 3). This length is consistent with the size of the longest cDNA, supporting the conclusion that we have identified a nearly full-length transcript. Importantly, this mRNA was completely lacking in mRNA prepared from *lsd1* leaves, furthering the argument that it encodes LSD1. The finding that *lsd1* is an mRNA allele was corroborated by sequencing across the genomic deletion in the mutant (Figure 3). The 5' border of the deletion is an A at nucleotide 55 and the 3' boundary is in the fourth intron (data not shown). It is noteworthy that expression of this candidate mRNA was unaffected by application of INA (Figure 4, top). The expected high level of INA-induced PR-1 mRNA accumulation in leaves of both wild type and *lsd1* (Figure 4, middle) served as a control in this experiment for efficacy of INA treatment.

The *lsd1* phenotype can be observed in all cell types examined after initiation of lesion formation (Dietrich et al., 1994). RNA blot analysis of seedlings, stems, leaves and flowers demonstrated that the *LSD1* gene is expressed constitutively in each of these Arabidopsis tissues (data not shown). Thus, the requirement for *LSD1* activity in these tissues is consistent with the gene's expression pattern.

Example XI **The *LSD1* mRNA encodes a novel zinc-finger domain**

We searched a variety of databases with the predicted translation product of the *LSD1* cDNA sequence. Several striking features emerged. First, there are three zinc-finger domains, depicted in Figure 5 (SEQ ID NOS:1-3), which share remarkable homology with one another. These are C-x-x-C, or type IV, zinc-fingers, according to the classification of Sánchez-Garía and Rabbitts (1994), and they share most homology with plant relatives of the GATA-1 transcription factor (Evans and Felsenfeld, 1989; Omichinski et al., 1993). The plant members of this sub-family described to date include the *CO* gene, which controls

transition to flowering (Putterill et al., 1995), a set of related DNA binding proteins (Yanagisawa, 1995; De Paolis et al., 1996) and a gene whose transcription is salt stress-induced (Lippuner et al., 1996). None of these proteins shares with LSD1 the consensus homology within the Zn-fingers. The second homology domain is derived from the carboxyl portion of LSD1, from residues 129 to 180 (Figure 6-SEQ ID NO:4). This region of LSD1 exhibits homology to three broad classes of regulatory proteins. First, all mammalian insulin receptor substrates; second, a set of animal transcription factors; and third, a maize transcription initiator binding protein.

The conceptual LSD1 translation product also identified two additional Arabidopsis ESTs via their predicted amino acid homology. Importantly, each has at least one C-x-x-C Zn-finger and most of the associated consensus residues found in the LSD1 internal homologies. They are ESTs 172A7T7 (GenBank R6552)(SEQ ID NO: 58 and 132J21T7 (GenBank T45809). Thus, it is probable that LSD1 is the first member of a widely distributed Zn-finger sub-family in plants, defined by the internal homology within each zinc-finger. The other amino acids in the consensus section are not known to be found in any other zinc finger proteins.

<u>Example XII</u> **Identification of expressed target sequence tags (EST) and cDNAs containing LSD1-type zinc finger domains**

As discussed in the text prior to the Examples, the predicted amino acid sequence of the LSD1 zinc fingers was used to search the GenBank database (NCBI). Two *Arabidopsis thaliana* ESTs (EST132J21T7 and EST 172A7T7) were identified, each of which contains at least two zinc finger domains and most of the associated consensus residues found in the LSD1 internal homologies (Dietrich, 1997). These ESTs were ordered from Ohio State University Arabidopsis Biological Resource Stock Center and resequenced. Sequences were analyzed with the Genetics Computer Group programs (Devereaux et al.,1994). A specific probe isolated from EST172A7T7 was subsequently used for screening of cDNA and genomic libraries. The bacterial strain carrying EST132J21T7, however, was not viable. Therefore, degenerated primers were designed based on the EST132J21T7 sequence. Genomic *Arabidopsis thaliana* Ws-0 DNA was used in the PCR reaction and gave rise to a specific PCR product of approximately 400 bp. This fragment was subcloned via the TA Cloning Kit (Invitrogen, Carlsbad, CA) into pBluescript KS(+). Two new genes were identified as described here. Their predicted protein products are highly related to that of LSD1 indicating an involvement in the control of cell death in plants

<u>Example XIII</u> ***LOL1* cDNA**

Poly A + RNA isolated from uninduced and *P. syringae* DC3000 induced

Arabidopsis thaliana Col-0 leaf tissue was reverse transcribed. The resulting cDNA population was subcloned unidirectionally into the EcoRI/XhoI – sites of a lambda-Zap II vector using the cDNA-synthesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's directions. The titer of this MG-library was calculated as 2.5×10^6 pfu Approximately 8×10^5 pfu of the amplified MG-library were subsequently screened with $\alpha^{32}P$ dCTP labeled probes (Stratagene 'Prime it' Kit) specific for EST132J21T7 or EST172A7T7. With the probe specific for EST132J21T7, four cDNA clones were identified and subcloned via the Stratagene excision system. One clone contained an insert of less than 100 bp in length and was not further analyzed. The three remaining clones were sequenced by standard protocol (primers: M13F, M13R, PE6, and PE7); for primer sequences refer to Table 5, below). Clones 2 and 3 contained identical open reading frames (ORFs) and were homologous to EST132J21T7 and to another identical and overlapping EST clone, EST119C9T7. The fourth clone consisted of a chimeric cDNA of approximately 1500 bp, with approximately 400 bp similarity to EST132J21T7, EST119C9T7, and clones 2 and 3. It was also not analyzed further.

Table 5. Primers and primer sequences used

Primer	Primer Sequence	SEQ ID NO:
M13F	5'- GTA AAA CGA CGG CCA TG -3'	37
M13R	5'- GGA AAC AGC TAT GAC CAT G -3'	38
PE6	5'- TTC ATG GCA ATG GTG TGA CCC C -3'	39
PE7	5'- CTG CCG GAT TCT TGA TCG AAG A -3'	40
PE8	5'- AGA GGA AGG TCC GCC TCC GG -3'	41
PE9	5'- CTC TGC TCT CCT GAG ACT GCT T -3'	42
PE13	5'- CAT CAT AAT GTC TCC TTT TGA GAC -3'	43
PE15	5'- GCC ATC CAT TAT TCA TCG CCT -3'	44
PE23	5'- GAG GAG GAA GAA CTG CAG ATT CC -3'	45
PE30	5'- GTG CTC CAT GTC CAA ATC ATA C -3'	46

Clone 2, with an insert length of 908 bp represents a full length cDNA clone, as determined by the presence of an open reading frame flanked by untranslated sequences, and was renamed *LOL1* (*L*sd *one like*)(SEQ ID NO:47). We confirmed that the *LOL1* cDNA and EST132J21T7 are encoded by the same gene using genomic DNA (Southern) blot analysis (data not shown). The LOL1 protein of 154 amino acids (SEQ ID NO:48) contains three zinc finger domains of the LSD1-type (SEQ ID NOS:49-51). The consensus

sequence of the LOL1 zinc finger domains is defined by CxxCxxLLMYxxGAxSxCxxC (SEQ ID NO:53).

Example XIV ***LOL2* cDNA**

By screening the MG-cDNA-library, no clones homologous to EST172A7T7 could be obtained. Therefore, the AB-cDNA-library (derived from RNA isolated from different tissues of sterile grown plants, available at the European Arabidopsis Stock Center, Cologne, Germany) was screened with $\alpha\ ^{32}P$ dCTP labeled probe specific for EST172A7T7. Six homologous cDNA clones were obtained and subcloned into the SmaI site of pBluescript KS(+). Restriction analysis indicated that the inserts were encoded by the same gene. Only the longest insert was sequenced following standard protocol (primers used: M13F, M13R, PE8 and PE9: for primer sequences, refer to Table 5. We demonstrated that this insert contained an ORF of 500 bp homologous to EST172A7T7. This non-full length cDNA was designated *LOL2* (SEQ ID NO:54). The deduced protein (SEQ ID NO:55) consisting of two LSD1-type zinc finger domains extending from bases 130-195 and 244-309 of SEQ ID NO:54 (SEQ ID NOS:56-57, respectively). Comparison to EST172A7T7 shows that the EST (SEQ ID NO:58) contains a 124 bp insertion (bases 386-509 after the second zinc finger of SEQ ID NO:58), leading to a different C-terminal. Comparison of these two partial cDNA sequences with the genomic *LOL2* sequence (see below) demonstrates that they are alternate splice forms from the same gene encoding two related proteins. This conclusion is strengthened by the fact that the *LOL2* cDNA and EST172A7T7 hybridize to the same genomic DNA fragment and therefore are encoded by the same gene (data not shown). Thus, sequence analysis of genomic *LOL2* clones shows that the non-identical C-termini of LOL2 and EST172A7T7 are due to alternative splice sites. The genomic sequence of *LOL2* (SEQ ID NO:59, has a putative TATA-box sequence and polyadenylation signal (bases 922-930 and 2539-2544), and the exon borders of an alternative splice site (bases 2256-2382). The derived amino acid sequence extends from bases 1231-2462.

Example XV **Isolation of genomic *LOL2* sequences from an *Arabidopsis thaliana* Col-0 library**

8×10^5 genomic lambda clones (lambda GEM11, European Arabidopsis Stock Center) were screen with a $\alpha\ ^{32}P$ dCTP labeled probe specific for EST172A7T7. Nine clones homologous to *LOL2* EST172A7T7 could be identified. Restriction analysis demonstrated that the nine clones belonged to five different classes. Inserts ranging from two to five kb in size were isolated and subcloned into either SacI ore BamHI sites of pBluescript KS(+). Sequence information derives from two overlapping clones,

sequentially sequenced with primers M13R, PE9, PE13, PE15, PE23 and PE30 (see Table 5).

The genomic *LOL2* sequence has a length of 3060 bp. Promoter and 5' untranslated regions consist of approximately 1200 bp. The translation products are encoded by three exons, which are interrupted by two introns of 182 bp and 458 bp length, respectively. The overall length of the coding sequence is 1232 bp. Due to alternative splice sites, two proteins which differ in their C-terminal regions are encoded by the *LOL2* gene (SEQ ID NO:59). A first protein, of 155 amino acids (SEQ ID NO:60), is identical to the *LOL2* cDNA and contains two zinc finger domains of the LSD1-type. The other translation product corresponds to EST172A7T7, consists of 147 amino acids, and contains two and a half zinc finger domains (SEQ ID NO:61). The consensus sequence of the two zinc finger domains of LOL2 is CxxCxxLLxYxxGxxxVxCSSC (SEQ ID NO:62).

<u>Example XVI</u> **Obtaining interacting genes**

The methodology for this Example is known to those skilled in the art and summarized in Fields and Sternglanz, 1994, and Finley and Brent, 1996. The *LSD1* short or *LDS1* long open reading frames were cloned into the "bait vector" pEG202 of the commonly available LexA yeast two-hybrid system (Matchmaker™, Clonetech, Palo Alto, CA) to generate plasmids pEG202-L and pEG202-S. These encode fusion proteins of the LexA DNA binding domain and the full length LSD1 protein of both long and short isoforms (SEQ ID NOS 14 and 15). Yeast strain EGY48 is transformed with this plasmid, and appropriate controls performed to ascertain the LSD1 fusion protein encoded by plasmids pEG202-L and pEG202-S did not intrinsically activate expression of the yeast markers used in this system. A yeast gene expression library was constructed in plasmid pJG4-5 using RNA from *Arabidopsis* leaves infected with *Pseudomonas syringae*. This library encodes fusion proteins of expressed *Arabidopsis* genes and the B42 transcriptional activation domain. The library was transformed en masse into the yeast strain EGY48 carrying either plasmid pEG202-L or -S. From an equivalent of 6 million clones screened, 122 were isolated. The longest insert of a member from each of these classes was sequenced using standard DNA sequencing methods. Because the novel *Arabidopsis* gene so identified is produced as an active translation fusion in this system, one is immediately able to identify the deduced protein sequence. The most interesting sequences thus defined, and their deduced protein sequences, are set forth herein as SEQ ID NOS: 66-91.

The first main class of LSD1-interacting proteins has no database homologues. These proteins encode putative "sequestration" proteins for LSD1 whose function is to inhibit LSD1 function until the correct pathogen signal is received. Their utility lies in

manipulation of the interaction with LSD1 in plant cells such that LSD1 is altered in its ability to regulate the response to pathogen. Alternatively, these novel LSD1-interacting proteins may encode new components of the gene regulation machinery working together with LSD1 to control transcription in response to pathogen infection. These proteins are valuable because of the knowledge that LSD1 is a key regulator of cell death in plants in response to pathogens. Proteins which physically interact with LSD1 share in this cellular function.

The second class defines proteins having database homologies to other proteins, strongly suggesting a role in control of gene transcription (e.g., CAAT box binding proteins which are known to bind the common CAAT regulatory unit in DNA preceding nearly all genes encoding eukaryotic mRNA). This finding is completely consistent with the embodiment described above, in which the LSD1 partner proteins identify other components of the gene regulatory machinery required for response to pathogens. Manipulation of the expression of, for example, CAAT box binding proteins, will result in altered response to pathogen infection.

While the invention has been described with reference to specific embodiments, it will be appreciated that numerous variations, modifications, and embodiments are possible, and accordingly, all such variations, modifications, and embodiments are to be regarded as being within the spirit and scope of the invention.

LITERATURE CITED

Baker, C. J. and Orlandi, E. W. (1995). Active oxygen in plant pathogenesis. Annu. Rev. Phytopathol. *33*, 299-322.

Bechtold, N., Ellis, J. and Pelletier, G. (1993). *In planta agrobacterium*-mediated gene transfer by infiltration of *Arabidopsis thaliana plants.* C. R. Acad. Sci., Paris *316*, 1194-1199.

Becker, D., Kemper, E., Schell, J. and Masterson, R. (1992). New plant binary vectors with selectable markers located proximal to the left T-DNA border. Plant Mol. Biol. *20*, 1195-1197.

Bent, A. (1996). Function meets structure in the study of plant disease resistance genes. Plant Cell *8*, 1757-1771.

Bolwell, G. P., Butt, V. S., Davies, D. R. and Zimmerlin, A. (1995). The origin of the oxidative burst in plants. Free Rad. Res. *23*, 517-532.

Brisson, L. F., Tenhaken, R. and Lamb, C. J. (1994). Function of oxidative cross-linking of cell wall structural proteins in plant disease resistance. Plant Cell *6*, 1703-1712.

Chen, Z., Silva, H. and Klessig, D. (1994). Involvement of reactive oxygen species in the induction of systemic acquired resistance by salicylic acid in plants. Science *262*, 1883-1886.

Choi, S., Creelman, R. A., Mullet, J. E. and Wing, R. A. (1995). Construction and characterization of a bacterial artificial chromosome library of *Arabidopsis thaliana*. Plant Mol. Biol. *13*, 124-128.

Church and Gilbert, (1984). Proc. Natl. Acad. Sci. USA 81:1991-1995.

Cosio, E. G., Frey, T. and Ebel, J. (1992). Identification of a high affinity binding protein for a hepta-β-glucoside phytoalexin elicitor in soybean. Eur. J. Biochem. *204*, 1115-1123.

Crute, I. R. (1985). The genetic bases of relationships between microbial parasites and their hosts. In Mechanisms of Resistance to Plant Diseases., R. S. S. Fraser, eds. (Dordrecht:

Kluwer Academic Press), pp. 80-143.

Dangl, J. L. (1995). Pièce de Résistance: Novel classes of plant disease resistance genes. Cell *80*, 363-366.

Dangl, J. L., Dietrich, R. A. and Richberg, M. H. (1996). Death Don't Have No Mercy: Cell Death Programs in Plant-Microbe Interactions. Plant Cell *8*, 1793-1807.

De Paolis, A., Sabatini, S., De Pascalis, L., Costantino, P. and Capone, I. (1996). A *rolB* regulatory factor belongs to a new class of single zinc finger plant proteins. Plant J. *10*, 215-224.

Delaney, T., Uknes, S., Vernooij, B., Friedrich, L., Weymann, K., Negrotto, D., Gaffney, T., Gut-Rella, M., Kessman, H., Ward, E. and Ryals, J. (1994). A central role of salicylic acid in plant disease resistance. Science *266*, 1247-1250.

Devereaux, J., Haeberli, M.J., and Smithies, O. (1984). Nucl. Acids Res. *12*, 387-395.

Dietrich, R. A., Delaney, T. P., Uknes, S. J., Ward, E. J., Ryals, J. A. and Dangl, J. L. (1994). Arabidopsis mutants simulating disease resistance response. Cell *77*, 565-578.

Dietrich, R.A., Richberg, M.H., Schmidt, R.., Dean, C., and Dangl, J.L. (1997). Cell *88*, 685-694.

Doke, N. (1983). Involvement of superoxide anion generation in the hypersensitive response of potato tuber tissues to infection with an incompatible race of *Phytophthora infestans* and to the hyphal wall components. Physiol. Plant Pathol. *23*, 345-357.

Durner, J. and Klessig, D. F. (1995). Inhibition of ascorbate peroxidase by salicylic acid and 2,6-dichloroisonicotinic acid, two inducers of plant defense responses. Proc. Natl. Acad. Sci., USA *92*, 11312-11316.

Enyedi, A. J., Yalpani, N., Silverman, P. and Raskin, I. (1992). Localization, conjugation and function of salicylic acid in tobacco during the hypersensitive reaction to tobacco mosaic virus. Proc. Natl. Acad. Sci., USA *89*, 2480-2484.

Evans, T. and Felsenfeld, G. (1989). The erytroid-specific transcription factor EryF1: a new

finger protein. Cell *58*, 877-885.

Feldmann, K. A. (1991). T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. Plant J. *1*, 71-82.

Fields, S. and Sternglanz, R. (1994). The two-hybrid system: an assay for protein-protein interactions. Trends Genet. *8*, 286-293.

Finley, R.L. and Brent, R. (1996). Interaction trap cloning in yeast. In Gene Probes: A Practical Approach, eds. Oxford: Oxford Univ. Press, pp._____

Flor, H. H. (1971). Current status of the gene-for-gene concept. Annu. Rev. Phytopathol. *9*, 275-296.

Gaffney, T., Friedrich, L., Vernooij, B., Negrotto, D., Nye, G., Uknes, S., Ward, E. and Ryals, J. (1993). Requirement for salicylic acid for the induction of systemic acquired resistance. Science *261*, 754-756.

Godiard, L., Grant, M. R., Dietrich, R. A., Kiedrowski, S. and Dangl, J. L. (1994). Perception and response in plant disease resistance. Curr. Opin. Genet. and Dev. *4*, 662-671.

Goodman, R. N. and Novacky, A. J. (1994). The hypersensitive response in plants to pathogens. (St. Paul: APS Press).

Gopalan, S., Bauer, D. W., Alfano, J. R., Lonlello, A. O., He, S. Y. and Collmer, A. (1996). Expression of the *Pseudomonas syringae* avirulence protein AvrB in plant cells alleviates its dependence on the hypersensitive response and pathogenicity (Hrp) secretion system in eliciting genotype-specific hypersensitive cell death. Plant Cell *8*, 1095-1105.

Grant, M. R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R. W. and Dangl, J. L. (1995). Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. Science *269*, 843-846.

Greenberg, J. T. and Ausubel, F. M. (1993). Arabidopsis mutants compromised for the control of cellular damage during pathogenesis and aging. Plant J. *4*, 327-342.

Greenberg, J. T., Guo, A., Klessig, D. F. and Ausubel, F. M. (1994). Programmed cell death

in plants: A pathogen-triggered response activated coordinately with multiple defense functions. Cell *77*, 551-564.

Hammond-Kosack, K. E. and Jones, J. D. G. (1996a). Inducible plant defense mechanisms and resistance gene function. Plant Cell *8*, in press.

Jabs, T., Dietrich, R. A. and Dangl, J. L. (1996). Initiation of runaway cell death in an Arabidopsis mutant by extracellular superoxide. Science *273*, 1853-1856.

Johal, G. S., Hulbert, S. H. and Briggs, S. P. (1994). Disease lesion mimics of maize: A model for cell death in plants. Bioessays *17*, 685-692.

Kauss, H. and Jeblick, W. (1995). Pretreatment of parsley suspension cultures with salicylic acid enhances spontaneous and elicited production of H2O2. Plant Physiol. *108*, 1171-1178.

Kauss, H., Theisinger-Hinkel, E., Mindermann, R. and Conrath, U. (1992). Dichloroisonicotinic and salicylic acid, inducers of systemic acquired resistance, enhance fungal elicitor responses in parsley cells. Plant J. *2*, 655-660.

Keen, N. T. (1990). Gene-for-gene complementarity in plant-pathogen interactions. Annu. Rev. Genet. *24*, 447-463.

Kiyosawa, S. (1970). Inheritance of a particular sensitivity of the rice variety Sekiguchi Asahi, to pathogens and chemicals, and linkage relationships with blast resistance genes. Bull. Nat. Inst. Agric. Sci. (Jpn) Ser D, Physiol. Genet. *21*, 61-71.

Kobayashi, D. Y., Tamaki, S. J. and Keen, N. T. (1989). Cloned avirulence genes from the tomato pathogen *Pseudomonas syringae* pv. *tomato* confer cultivar specificity on soybean. Proc. Natl. Acad. Sci., USA *86*, 157-161.

Konieczny, A. and Ausubel, F. M. (1993). A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant J. *4*, 403-410.

Kosslak, R. M., Chamberlin, M. A., Palmer, R. G. and Bowen, B. A. (1996). Apoptosis and necrosis in adjacent root cortical cells precede the defense response in soybean *ROOT NECROSIS* mutants. J. Hered. 87, 413-422.

Langford, A. N. (1948). Autogenous necrosis in tomatoes immune from *Cladosporium fulvum* Cooke. Can. J. Res. *26*, 35-64.

Leon, J., Lawton, M. A. and Raskin, I. (1995). Hydrogen peroxide stimulates salicylic acid biosynthesis in tobacco. Plant Physiol. *108*, 1673-1678.

Levine, A., Pennell, R., Palmer, R. and Lamb, C. J. (1996). Calcium-mediated apoptosis in a plant hypersensitive response. Curr. Biol. *6*, 427-437.

Levine, A., Tenhaken, R., Dixon, R. and Lamb, C. J. (1994). H2O2 from the oxidative burst orchestrates the plant hypersensitive disease resistance response. Cell *79*, 583-593.

Lippuner, V., Cyert, M. S. and Gasser, C. S. (1996). Two classes of plant cDNA clones differentially complement yeast calcineurin mutants and increase salt tolerance of wild-type yeast. J. Biol. Chem. *271*, 12859-12866.

Lister, C. and Dean, C. (1993). Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. Plant J. *4*, 745-750..

Low, P. S. and Merida, J. R. (1996). The oxidative burst in plant defense: Function and signal transduction. Physiol. Plant. *96*, 533-542.

Lukowitz, W., Mayer, U. and Jürgens, G. (1996). Cytokinesis in the Arabidopsis embryo involves the syntaxin-related KNOLLE gene product. Cell *84*, 61-71.

Malamy, J., Hennig, J. and Klessig, D. F. (1992). Temperature-dependent induction of Salicylic acid and its conjugates during the resistance response to tobacco mosaic virus infection. The Plant Cell *4*, 359-366.

Matallana, E., Bell, C. J., Lu, P. J. and Ecker, J. E. (1992). Genetic and physical linkage of the Arabidopsis genome: Methods for anchoring yeast artificial chromosomes. In Methods in Arabidopsis Research, C. Koncz, N.-H. Chua and J. Schell, eds. (Singapore: World Scientifc), pp. 144-169.

Mauch-Mani, B. and Slusarenko, A. J. (1996). Production of salicylic acid precursors is a major function of phenylalanine-ammonia lyase in the resistance of Arabidopsis to *Peronospora parasitica*. Plant Cell *9*, 203-212.

May, M. J., Hammond-Kosack, K. E. and Jones, J. D. G. (1996). Involvement of reactive oxygen species, glutathione metabolism and lipid peroxidation in the *Cf*-gene-dependent defense response of tomato cotyledons induced by race-specific elicitors of *Cladosporium fulvum*. Plant Physiol. *110*, 1367-1380.

Mittler, R., Shulaev, V. and Lam, E. (1995). Coordinated activation of programmed cell death and defense mechanisms in transgenic tobacco plants expressing a bacterial proton pump. Plant Cell *7*, 29-42.

Newman, T., de Bruijn, F. J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C. (1994). Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. Plant Physiol. *106*, 1241-1255.

Nürnberger, T., Nennstiel, D., Jabs, T., Sacks, W. R., Hahlbrock, K. and Scheel, D. (1994). High-affinity binding of a fungal oligopeptide elicitor to parsley plasma membranes triggers multiple defense responses. Cell *78*, 449-460.

Omichinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J. and Gronenborn, A. M. (1993). NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. Science *261*, 438-446.

Pryor, T. P. and Ellis, J. (1993). The genetic complexity of fungal resistance genes in plants. Adv. Plant Pathol. *10*, 281-307.

Putterill, J., Robson, F., Lee, K., Simon, R. and Coupland, G. (1995). The *CONSTANS* gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell *80*, 847-857.

Ramain, P., Heitzler, P., Haenlin, M. and Simpson, P. (1993). *pannier*, a negative regulator of *achaete* and *scute* in *Drosophila*, encodes a zinc finger protein with homology to the vertebrate transcription factor GATA-1. Development *119*, 1277-1291.

Rogers, S. O. and Bendich, A. J. (1985). Extraction of DNA from milligram quantities of fresh, herbarium, and mummified plant tissue. Plant Mol. Biol. *5*, 69-76.

Rounsley, S. D., Glodek, A., Sutton, G., Adams, M. D., Somerville, C. R., Venter, J. C. and Kerlavage, A. R. (1996). The construction of Arabidopsis expressed sequence tag assemblies. A new resource to facilitate gene identification. Plant Physiol. *112*, 1177-1183.

Rueffler, M., Steipe, B. and Zenk, M. H. (1995). Evidence against specific binding of salicylic acid to plant catalase. FEBS Letters. *377*, 175-180.

Ryals, J. L., Neuenschwander, U. H., Willits, M. C., Molina, A., Steiner, H.-Y. and Hunt, M. D. (1996). Systemic acquired resistance. Plant Cell *8*, 1809-1819.

Ryerson, D. E. and Heath, M. C. (1996). Cleavage of nuclear DNA into oligonucleosomal fragments during cell death induced by fungal infection or by abiotic treatments. Plant Cell *8*, 393-402.

Sánchez-Garía, I. and Rabbitts, T. H. (1994). The LIM motif: a new structural motif found in zinc-finger-like proteins. Trends Genet. *10*, 315-320.

Schmidt, R., West, J., Cnops, G., Love, K., Balestrazzi, A. and Dean, C. (1996). Detailed description of four YAC contigs representing 17Mb of chromosome 4 of *Arabidopsis thaliana* ecotype Columbia. Plant J. *9*, 755-765.

Schmidt, R., West, J., Love, K., Lenehan, Z., Lister, C., Thompson, H., Bouchez, D. and Dean, C. (1995). Physical map and organization of *Arabidopsis thaliana* chromosome 4. Science *270*, 480-483.

Scofield, S. R., Tobias, C. M., Rathjen, J. P., Chang, J. H., Lavelle, D. T., Michelmore, R. W. and Staskawicz, B. J. (1996). Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. Science *in press*,

Stakman, E. C. (1915). Relation between *Puccinia graminis* and plants highly resistant to its attack. J. Agric. Res. *4*, 193-299.

Staskawicz, B. J., Ausubel, F. M., Baker, B. J., Ellis, J. and Jones, J. D. G. (1995). Molecular genetics of plant disease resistance. Science *268*, 661-667.

Tang, X., Frederick, R. D., Zhou, J., Halterman, D. A., Jia, Y. and Martin, G. B. (1996). Physical interaction of avrPto and the Pto kinase defines a recognition event involved in

plant disease resistance. Science *in press*,

Uknes, S., Dincher, S., Friedrich, L., Negrotto, D., Williams, S., Thompson-Taylor, H., Potter, S., Ward, E. and Ryals, J. (1993a). Regulation of pathogenesis-related protein 1-a gene expression in tobacco. Plant Cell *5*, 159-169.

Uknes, S., Winter, A. M., Delaney, T., Vernooij, B., Morse, A., Friedrich, L., Nye, G., Potter, S., Ward, E. and Ryals, J. (1993b). Biological induction of systemic acquired resistance in *Arabidopsis*. Molec. Plant-Microbe Interact. *6*, 692-698.

Valent, B., Farrall, L. and Chumley, F. G. (1990). *Magnaporthe grisea* genes for pathogenicity and virulence identified through a series of backcrosses. Genetics *127*, 87-101.

Vernooij, B., Friedrich, L., Morse, A., Reist, R., Kolditz-Jawhar, R., Ward, E., Uknes, S., Kessmann, H. and Ryals, J. (1994). Salicylic acid is not the translocated signal responsible for inducing systemic acquired resistance, but is required in signal transduction. Plant Cell *6*, 959-965.

Walbot, V., Hoisington, D. A. and Neuffer, M. G. (1983). Disease lesion mimics in maize. In Genetic Engineering of Plants, T. Kosuge and C. Meredith, eds. (New York: Plenum Publishing Co.), pp. 431-442.

Wang, C.-Y., Mayo, M. W. and Baldwin Jr., A. S. (1996a). TNF- and cancer therapy-induced apoptosis: potentiation by inhibition of NF-κB. Science *274*, 784-787.

Wang, H., Li, J., Bostock, R. M. and Gilchrist, D. G. (1996b). Apoptosis: A functional paradigm for programmed plant cell death induced by a host-selective phytotoxin and invoked during development. Plant Cell *8*, 375-391.

Whalen, M. C., Stall, R. E. and Staskawicz, B. J. (1988). Characterization of a gene from a tomato pathogen determining hypersensitive resistance in non-host species and genetic analysis of this resistance in bean. Proc. Natl. Acad. Sci., USA *85*, 6743-6747.

Yanagisawa, S. (1995). A novel DNA-binding domain that may form a single zinc finger. Nucl. Acids Res. *23*, 3403-3410.

I

SEQUENCE LISTING

SEQ ID NO:1

LVCHGCRNLLMYPRGASNVRCALCNTINMV

SEQ ID NO:2

IICGGCRTMLMYTRGASSVRCSCCQTTNLV

SEQ ID NO:3

INCGHCRTTLMYPYGASSVKCAVCQFVTNV

SEQ ID NO:4

MSNGRV-PLPTNRP-NGTACPPST-STSTPPSQTQTVVVENPMSVDESGKLVSNV

SEQ ID NO:5

MSPG-VAPVPSNRKGNGDYMPMSPKSVSAP-QQIINPIRRHPQRVDPNGYMM

SEQ ID NO:6

VPLPTNRP-NGTACPPSTSTSTPPSQTQTVVVENPMSVDESGKLVSNV

SEQ ID NO:7

VPLPANNPVV-TTVVPSTPPSQPPAVCPPVV

SEQ ID NO:8

VPLPANNPVV-TTVVPSTPPSQPPAVCPPVV

SEQ ID NO:9

IPVYTNSNV-GTALPPSVSPSVSPSVT

SEQ ID NO:10

VVLP-NAAPAGAAAPPSGSRSTSPS

SEQ ID NO:11

SNGRVPLPTNRPN-GTACPPSTSTSTPPSQTQTVVVENPMSVDESGKLVSNV

SEQ ID NO:12

SRALVPVPAADPNAG-AIVPANKSKRSPEQGQRRIRR

SEQ ID NO:13

```
        10                            30                            50
GATCAAATCTAGTTACGCTTAAATTTGGATATATCTAAGGTTTCTTCGTCAATATATGGA

        70                            90                           110
GCTTACGAAAACGAAAGAGTGAGCTACGAGGAACTAAATCAATGAAGATAAGAGGAATGA

       130                           150                           170
AGGAGAGAAGATCACCAAGGTGTAGAAATTTCTGAAGTCGTCTCCTCCAATCTCCACTAT

       190                           210                           230
TGGTTTGTTCAGAACTTGAGAAGGCCTTAGATCCAAGCCATTAGTAACCTCTCTATGGCC

       250                           270                           290
ATAAGTGACCTTAAGAGAGACCAACCTCGTGAAAGGATCAAGAACATCTCCAACAACACT

       310                           330                           350
GCCGACCACGAGAGGATCTCTACGACTTAAAGACATATTTATCTTGGATCTCAAGTATCT

       370                           390                           410
CAATAAAATGTTTTGCTTCTAACCTTATGAACCCTTACTTGCTATTCTTTATATAACGTT

       430                           450                           470
TTGGGAATTGCAATAATTAGCTATTTAGCTTTATTCTCTCCAATGAAATCATTACCAGGG

       490                           510                           530
TCTTTTCGTGTATAGTTATCTTCGAGAATCTACAACTCGTTCAACGTACGTATATCACTT

       550                           570                           590
ATAATTCATGTTTTTTTTTCTTTCCTTTTTTCTAAATTTATAGTATTCTTATTCCAAAAC

       610                           630                           650
CCACCAGTATAAAACAGAAATAATCATATTCCAAATTATACATCATCCACTTGTTTCTTG

       670                           690                           710
CTAGCCACTAGTATGTAATTTATTCTGACTTATCATTGGAACTTCATGAACTATTTAAAA

       730                           750                           770
TAATGTCACAAGCATATAATATGCTGCATATTTGCGTACGTCACGCATTTTGCGTCACAT

       790                           810                           830
GTCACTCATTTAAATAGTTAAGGACACTACATTACACCGATTATGTATGATGTTAATGCA

       850                           870                           890
TTTTAGAATAACTCCTTCAACCTAAACCATCATATAAAAGTATATATGCTCCAGATAAAT

       910                           930                           950
TGACGCCATAATTGTTCACATATCTGGTTGGTTTGTACATACGTACTAGACTCTTTTTTT

       970                           990                          1010
TCTTTTCTTAATGTAGTACTAAACTTAATTAATACCATCAAAAATATCAATTTAACAAAA

      1030                          1050                          1070
```

3

```
CAAACCAGTAAAACTTTTAAAACAATGGAGTAAATCAAATAAAACAAGTAAATTAACAAA

        1090                1110                1130
TAGACACAAGGTAACAGAAGTATAATAACGACAGAAAAATGAACAATTGGCCAAAAAATT

        1150                1170                1190
CGTTTTCAAACGTGATTTCAAAATTGTCTCCAAATCTTAAATGTTGATAAAGTAATTTTT

        1210                1230                1250
TTTTAAATTCATTATACCTTTCAAAAACAAGTGTATTACCTAAAAGCTCAACCGTGTATT

        1270                1290                1310
CTTACACTCCAAACAAATTTAGTTCCCCAAGTTTGGAAGACAAAAATTTCTAAGAAATTT

        1330                1350                1370
CTGACAAAACACATGAGAAATAAACCGATAAAGACTTCTAAAAACTATTGCAGACCAGTT

        1390                1410                1430
TCATTTGCTGACCACAAAAAGTCATGAGAATACAATTAGCTCAGTGATTCTTGATATTTC

        1450                1470                1490
TGGTACCTAACAAAAAGAAAAGTGTGTGAGGTTAGATGGCTATGATTTTTGCTCTCCAAT

        1510                1530                1550
TTATTGTCCATTTCCCCAATTTGTAATATGAAATGCGCAAATTACTCTTCTTCCGATATG

        1570                1590                1610
AATAAGCAAACGAAAACATACGTGGGACGTTATGTTGAGAACATTTGATTAAAGTTTATA

        1630                1650                1670
TGCGATTTTCATTTATTTACTATGAATTTTTTGTTTGGCAGCATGTACGATTTTTCATTT

        1690                1710                1730
AACACACAAATATTATAGAATTTTCATTGGTTCAAAGGGGTAGACAAAAAATAATTTAAT

        1750                1770                1790
ATTATTACACCATTTGCAGAAAATTAGAAAATATATTTTTACCCATAATTAATTGATCTA

        1810                1830                1850
TGGACGTATGCTTGGCATAAAAATTCATATTTAATTAGCAGAAGCCAATCGCTGCGTTTG

        1870                1890                1910
TATATACGCGTTTATGACCGAGAAAAAAACCCTTACGCGTCATGTAAAAAAAAAAAGAAGC

        1930                1950                1970
GTAAATTACGAAAAACAGAGAGATAAATCCGGGCATTGAGATTTTGGAGATAGAGAGAGA

        1990                2010                2030
GAAAAATCGAAATCTATTGTCTATCTCCTCAATTTGGATTGGATTTTCTGCATATCATCG

        2050                2070                2090
CTCTAGATTTCGCGGGTTTTGGATTCGATTCCTTACCCTTCTCCAATCGGTAAGAACAAG

        2110                2130                2150
CTCCAAAGTTTGTTCCTTTTTTTCAATTTTCGCCAATTCTGTAATCTCATCATTGTTCTT
```

4

2170 2190 2210
GTTTGATTTGGATGCAGAAGTTTTTGGGTTTGAATTGGATTTGGGTTTCGTTCCAAAATC

2230 2250 2270
AGCTCTTTTTGTTAATCAGGTGAGTTTTTAGGTATTTGAATCTCCAATTGCTTCCCTTGC

2290 2310 2330
AATGACTAAGTATTGTGAAATGTTTAGGGTTTCATCTGTGTGGGTCTTGTTTTGAAGCAA

2350 2370 2390
TTTGTGTGTGTTTGGATGAAAGTAGCAGATATGCAGGACCAGCTGGTGTGTCATGGTTGT

2410 2430 2450
AGGAATTTATTGATGTATCCTAGAGGAGCATCTAATGTGCGTTGTGCGTTATGTAACACT

2470 2490 2510
ATCAACATGGTTCCTCCTCCTCCTCCACCTCACGGTATCGATTTCTTTGTTGAATTTGAA

2530 2550 2570
TTGAGGATGAGGTTAATATGCTCTGCAATTGTATTATAACTTGGGTTCTGATTCTGAATA

2590 2610 2630
CAGACATGGCACACATTATATGTGGTGGTTGTAGAACAATGCTTATGTATACGCGTGGGG

2650 2670 2690
CTAGTAGCGTAAGATGCTCTTGCTGTCAAACTACGAACCTTGTGCCAGGTATATTAATAA

2710 2730 2750
TATCGTGACATCCATATCAATCCTTTTAAAGACCATGTATTATATTGCTTTATAAGGTCT

2770 2790 2810
TTTAGTCCTTTAGAATCTTCTTTCACACTTTTGTTTGATAACATTGTTCTGTGGAGATGA

2830 2850 2870
TGCTTACGTAACGTATTTCCACTTTTCCCAAAGATGTATATGAATCTGAATTCTGAAAAT

2890 2910 2930
ATCTGGGATTTGTAAAGCAGCTGAAAGTACTTAAAACAAAGCTTTTAGATGGTCCCGGTG

2950 2970 2990
GACTAGGTAACTACTTGTTAGAGCTAGTAGGGTTTATTATTGTTTTGTTTGATCTACCAT

3010 3030 3050
TAGATTCTTATCTTTAATTAGCGTCTAAGCTGTTGTCATTTAGCTGTATGATTATCATTT

3070 3090 3110
ATCCATGACTGCTTAAGAACATTGCTGATTACTTCGTTCATTAGTATTTCTTGGATTTTT

3130 3150 3170
CTAGCATTAACATTGCTTGTTTTCTGAATCTGTGCGTGTCTTTTTTGAAATCGACAGCGC

3190 3210 3230
ACTCCAATCAGGTTGCCCATGCTCCTTCCAGTCAGGTTGCGCAAATCAATTGTGGGCATT

3250 3270 3290
GTCGGACGACCCTCATGTATCCTTACGGTGCATCATCCGTCAAATGCGCTGTTTGTCAAT

3310 3330 3350
TCGTAACTAACGTTAATGTGATTATTCCTATCTATTAAGCCACCTCTGCATGGTTGAGTT

3370 3390 3410
AAGTATAGAGATCTTTCTGTTGGAAATTTTCATTTCTGATTCATTTTGCATCCTTAGATG

3430 3450 3470
AGCAATGGAAGGGGTACCTCTCCCAACTAACCGGCCAAATGGAACAGCTTGTCCCCCCTC

3490 3510 3530
TACATCAACTGTGAGTTATCAAATTATGAATTTGTAATAGTTCTGTATATTCTTATGGAA

3550 3570 3590
CTGGTACTTACTCTGTTCATCGATTTTTCATTTTACCAACAGTCAACACCACCCTCTCAG

3610 3630 3650
ACCCAAACCGTTGTTGTAGAAAACCCCATGTCCGTTGATGAAAGCGGAAAGTTGGTGAGT

3670 3690 3710
ATTTCTATCACCTGTGTTCTTCTTCTTATTTACCACATTAGAGGAAGATATGACAAAGTG

3730 3750 3770
ACTGAAACACACAAATTGCAGGTGAGCAATGTTGTTGTTGGAGTGACAACTGACAAAAAG

3790 3810 3830
TAATCAAGAATGAGTGAGATCTTAAAGATCAAATCCAAATTCTTCCTCTATTCCTGCGTT

3850 3870 3890
TGGTTTGTGCATATTACATACGCGGAAAAACTGTATGTTATATATCTCTTGACTCCTTTT

3910 3930 3950
TAACCCAAGAGAAAAAGCTTATCAGAATCTCTTGTTACTGCATTATTGGGGTTTATTCAA

3970 3990 4010
AGTTGAAGACACAAGGTTTTTGCTCGAATAATTTGGCATTCTTTTGCTCCATGGAACTTG

4030 4050 4070
ACCTTCTCTTCTGTTAGTTGACTTCTAAAACTCCATCGGCCCTTGTGGCATTGTTAATGT

4090 4110 4130
ATGTATGAATATAATCTGATACACCAACCAATCATTAAGATTTGGGTTTGAAATCTGTCT

4150 4170 4190
CTTCCGTGGATGAGATATGCTACATGTCACAAGAACTGGTCTTAGCTTTGGTAGATAAGA

4210 4230 4250
CTTGTCTTAGAAGCAAGTCTTGAAATCTGGAAATCTATTTTGCAGTAATCTTGTCACAAC

4270 4290 4310
AACCATAACCTAATCAGTCAGTACCCTCCAAGAAACATTAAAGTTAGATGATCCGACAAA

4330 4350 4370

6

ACCTCTCAACAAAACCAACTCTTTCCATATAAATACTCTTTAACACTGGACCAAATTTNC

4390 4410 4430
ACCCTTCCTCTTGATCCTCCCTGCATCACAATGGCCAAAAAAAAAATGGTGGTTGGCNGG

4450 4470 4490
TGGGTACCACAAAGAGCTGGAAACTACTCTTGGGGCTGAGAATATTTGCATTCATGGCTA

4510
CTTTAGCTGCAG

SEQ ID NO:14

10 30 50
CTTACGCGTCATGTAAAAAAAAAAGAAGCGTAAATTACGAAAAACAGAGAGATAAATCCG

70 90 110
GGCATTGAGATTTTGGAGATAGAGAGAGAGAAAAATCGAAATCTATTGTCTATCTCCTCA

130 150 170
ATTTGGATTGGATTTTCTGCATATCATCGCTCTAGATTTCGCGGGTTTTGGATTCGATTC

190 210 230
CTTACCCTTCTCCAATCGAAGTTTTTGGCTTTGAATTGGATTTGGGTTTCGTTCCAAAAT

250 270 290
CAGCTCTTTTTGTTAATCAGATATGCAGGACCAGCTGGTGTGTCATGGTTGTAGGAATTT

310 330 350
ATTGATGTATCCTAGAGGAGCATCTAATGTGCGTTGTGCGTTATGTAACACTATCAACAT

370 390 410
GGTTCCTCCTCCTCCTCCACCTCACGACATGGCACACATTATATGTGGTGGTTGTAGAAC

430 450 470
GATGCTTATGTATACGCGTGGGGCTAGTAGCGTAAGATGTTCTTGCTGTCAAACTACGAA

490 510 530
CCTTGTGCCAGCGCACTCCAATCAGGTTGCCCATGCTCCTTCCAGTCAGGTTGCGCAGAT

550 570 590
CAATTGTGGGCATTGTCGGACGACCCTCATGTATCCTTACGGTGCATCATCCGTCAAATG

610 630 650
CGCTGTTTGTCAATTCGTAACTAACGTTAATATGAGCAATGGAAGGGTACCTCTCCCAAC

670 690 710
TAACCGGCCAAATGGAACAGCTTGTCCCCCCTCTACATCAACTTCAACACCACCCTCTCA

730 750 770
GACCCAAACCGTTGTTGTAGAAAACCCCATGTCCGTTGATGAAAGCGGAAAGTTGGTGAG

790 810 830
CAATGTTGTTGTTGGAGTGACAACTGACAAAAAGTAATCAAGAATGAGTGAGATCTTAAA

7

```
     850                 870                 890
GATCAAATCCAAATTCTTCCTCTGTTCCTGCGTTTGGTTTGTGCATATTACATACGCGGA

     910                 930                 950
AAAACTGTATGTTATATATCTCTTGACTCCTTTTTAACCCAAGAGAAAAAGCTTATCAGA

     970
AAAAAAAAAAAAAAAAA
```

SEQ ID NO:15

```
      10                  30                  50
GAAATCTATTGTCTATCTCCTCAATTTGGATTGGATTTTCTGCATATCATCGCTCTAGCT

      70                  90                 110
TTCGCGGGTTTTGGATTCGATTCCTTACCCTTCTCCAATCGAAGTTTTTGGCTTTGAATT

     130                 150                 170
GGATTTGGGTTTCGTTCCAAAATCAGCTCTTTTTGTTAATCAGGGTTTCATCTGTGTGGG

     190                 210                 230
TCTTGTTTTGAAGCAATTTGTGTGTGTTTGGATGAAAGTAGCAGATATGCAGGACCAGCT

     250                 270                 290
GGTGTGTCATGGTTGTAGGAATTTATTGATGTATCCTAGAGGAGCATCTAATGTGCGTTG

     310                 330                 350
TGCGTTATGTAACACTATCAACATGGTTCCTCCTCCTCCTCCACCTCACGACATGGCACA

     370                 390                 410
CATTATATGTGGTGGTTGTAGAACAATGCTTATGTATACGCGTGGGGCTAGTAGCGTAAG

     430                 450                 470
ATGCTCTTGCTGTCAAACTACGAACCTTGTGCCAGCGCACTCCAATCAGGTTGCCCATGC

     490                 510                 530
TCCTTCCAGTCAGGTTGCGCAGATCAATTGTGGGCATTGTCGGACGACCCTCATGTATCC

     550                 570                 590
TTACGGTGCATCATCCGTCAAATGCGCTGTTTGTCAATTCGTAACTAACGTTAATATGAG

     610                 630                 650
CAATGGAAGGGTACCTCTCCCAACTAACCGGCCAAATGGAACAGCTTGTCCCCCCTCTAC

     670                 690                 710
ATCAACTTCAACACCACCCTCTCAGACCCAAACCGTTGTTGTAGAAAACCCCATGTCCGT

     730                 750                 770
TGATGAAAGCGGAAAGTTGGTGAGCAATGTTGTTGTTGGAGTGACAACTGACAAAAAGTA

     790                 810                 830
ATCAAGAATGAGTGAGATCTTAAAGATCAAATCCAAATTCTTCCTCTATTCCTGCGTTTG

     850                 870                 890
GTTTGTGCATATTACATACGCGGAAAAACTGTATGTTATATATCTCTTGACTCCTTTTTA
```

8

```
      910                 930                 950
ACCCAAGAGAAAAAGCTTATCAGAATCTCTTGTTACTGCATTATTGGGGTTTATTCAAAG

      970                 990
TTGAAGACACAAGGTTTTTGCTCGAAAAAAAAAAAAAAAAAAAAA
```

SEQ ID NO:16

```
MetGlnAspGlnLeuValCysHisGlyCysArgAsnLeuLeuMetTyrProArgGlyAla
                              10                                  20

SerAsnValArgCysAlaLeuCysAsnThrIleAsnMetValProProProProProPro
                              30                                  40

HisAspMetAlaHisIleIleCysGlyGlyCysArgThrMetLeuMetTyrThrArgGly
                              50                                  60

AlaSerSerValArgCysSerCysCysGlnThrThrAsnLeuValProAlaHisSerAsn
                              70                                  80

GlnValAlaHisAlaProSerSerGlnValAlaGlnIleAsnCysGlyHisCysArgThr
                              90                                 100

ThrLeuMetTyrProTyrGlyAlaSerSerValLysCysAlaValCysGlnPheValThr
                             110                                 120

AsnValAsnMetSerAsnGlyArgValProLeuProThrAsnArgProAsnGlyThrAla
                             130                                 140

CysProProSerThrSerThrSerThrProProSerGlnThrGlnThrValValValGlu
                             150                                 160

AsnProMetSerValAspGluSerGlyLysLeuValSerAsnValValValGlyValThr
                             170                                 180

ThrAspLysLys
```

SEQ ID NO:17

```
MetLysValAlaAspMetGlnAspGlnLeuValCysHisGlyCysArgAsnLeuLeuMet
                              10                                  20

TyrProArgGlyAlaSerAsnValArgCysAlaLeuCysAsnThrIleAsnMetValPro
                              30                                  40

ProProProProProHisAspMetAlaHisIleIleCysGlyGlyCysArgThrMetLeu
                              50                                  60

MetTyrThrArgGlyAlaSerSerValArgCysSerCysCysGlnThrThrAsnLeuVal
```

9

70 80

ProAlaHisSerAsnGlnValAlaHisAlaProSerSerGlnValAlaGlnIleAsnCys
90 100

GlyHisCysArgThrThrLeuMetTyrProTyrGlyAlaSerSerValLysCysAlaVal
110 120

CysGlnPheValThrAsnValAsnMetSerAsnGlyArgValProLeuProThrAsnArg
130 140

ProAsnGlyThrAlaCysProProSerThrSerThrSerThrProProSerGlnThrGln
150 160

ThrValValValGluAsnProMetSerValAspGluSerGlyLysLeuValSerAsnVal
170 180

ValValGlyValThrThrAspLysLys

SEQ ID NO:18

5'-CAG TGG ATC TTT CCT CAG ACG-3'

SEQ ID NO:19

5'-CAT CTT CTT CTG CAA TCT GGG-3'

SEQ ID NO:20

5'-CAT CCA TCA AAC AAA CTC C-3'

SEQ ID NO:21
5'-TGT TTC AGA GTA GCC AAT TC-3'

SEQ ID NO:22
5'-CAC GTT AGT TAG TTA GAA GG-3'

SEQ ID NO:23
5'-CTG ATG TTC TCT ACA AAT GG-3'

SEQ ID NO:24

5'-CGT ATC CGC ATT TCT TCA CTG C-3'

SEQ ID NO:25
5'-CAT CTG CAA CAT CTT CCC CAG-3'

SEQ ID NO:26
5'-TTG AGT CCT TCT TGT CTG-3'

SEQ ID NO:27
5'-CTA GAG CTT GAA AGT TGA TG-3'

SEQ ID NO:28
5'-GAA TGG TGT AAC CAA ACT C-3'

SEQ ID NO:29
5'-CAT ACC GTA TGA TGG AAC-3'

SEQ ID NO:30
5'-GAA CTC ATT GTA TGG ACC-3'

SEQ ID NO:31
5'-CTA AGA TGG GAA TGT TGG-3'

SEQ ID NO:32
5'-CCA AGA AGA GAA AAC GGA GA-3'

SEQ ID NO:33
5'-AAC AAT AGG AGG TGC AGA GT-3'

SEQ ID NO:34
5'-ACC TAA CAA AAA GAA AAG TGT GTG AGG-3'

SEQ ID NO:35
5'-ATA ATA AAC CCT ACT AGC TCT AAC AAG-3'

SEQ ID NO:36
5'-CTG CTA CTT TCA TCC AAA C-3'

11

SEQ ID NO:37
5'- GTA AAA CGA CGG CCA TG -3'

SEQ ID NO:38
5'- GGA AAC AGC TAT GAC CAT G -3'

SEQ ID NO:39
5'- TTC ATG GCA ATG GTG TGA CCC C -3'

SEQ ID NO:40
5'- CTG CCG GAT TCT TGA TCG AAG A -3'

SEQ ID NO:41
5'- AGA GGA AGG TCC GCC TCC GG -3'

SEQ ID NO:42
5'- CTC TGC TCT CCT GAG ACT GCT T -3'

SEQ ID NO:43
5'- CAT CAT AAT GTC TCC TTT TGA GAC -3'

SEQ ID NO:44
5'- GCC ATC CAT TAT TCA TCG CCT -3'

SEQ ID NO:45
5'- GAG GAG GAA GAA CTG CAG ATT CC -3'

SEQ ID NO:46
5'- GTG CTC CAT GTC CAA ATC ATA C -3'

SEQ ID NO:47

```
         10                  30                  50
AATATATCGAAACGAGATTCCACAATTAGTCTCTAGTCAAAGAGCTTCATGGCAATGGTG

         70                  90                  110
```

TGACCCCAAATATAGATTTGATGAAAGTGAGGAAATAGGAGAAGAAATGAAGAACACAGG

130 150 170
ATGTGTCTTCTTCTTCTAAGTCACTAACAAAATCAACAAAGAGGAGAAGCCATTATTATA

190 210 230
TAATAGAGAGATTGAGAGAAGAGATTTATCCAAAAAAATATTGCAATTCTTCTTGGAGTG

250 270 290
AATAATGCCAGTCCCTCTTGCACCATATCCAACACCTCCGGCACCGGCACAGGCTCCGTC

310 330 350
GTACAACACTCCTCCGGCAAATGGAAGTACAAGTGGGCAGAGCCAGTTAGTGTGTTCAGG

370 390 410
TTGCAGAAACCTTCTGATGTATCCCGTCGGAGCAACCTCCGTCTGCTGCGCCGTCTGTAA

430 450 470
CGCCGTCACGGCCGTTCCTCCGCCGGGAACGGAGATGGCACAGTTAGTATGTGGAGGATG

490 510 530
CCATACACTCTTAATGTACATTCGTGGAGCTACAAGTGTTCAATGTTCTTGTTGTCACAC

550 570 590
TGTTAATCTCGCCCTCGAAGCGAACCAAGTAGCGCATGTGAATTGCGGAAACTGCATGAT

610 630 650
GCTACTAATGTATCAATATGGAGCAAGATCAGTGAAATGTGCCGTTTGTAACTTTGTCAC

670 690 710
ATCTGTTGGGGGTTCAACGAGCACGACTGATTCGAAGTTTAACAATTAAAACTTGGATCT

730 750 770
ATCTACCTATCAATATCTATTGAGTTATGAGCAATATAGAGGAAGCATCAAATCTTTTTC

790 810 830
ACTCTCTCTTCGATCAAGAATCCGGCAGTTATGAGTTTGAAACCATTTTCGGAAGTAAAT

850 870 890
GAAATATGTAATTCGTCGAAATTTCTGACTTTGGTCTCTTTGTCCGTTTGTATAGAGCTA

910
AAAAAAAAAA

SEQ ID NO: 48

MetProValProLeuAlaProTyrProThrProProAlaProAlaGlnAlaProSerTyr
10 20

AsnThrProProAlaAsnGlySerThrSerGlyGlnSerGlnLeuValCysSerGlyCys

13

30 40

ArgAsnLeuLeuMetTyrProValGlyAlaThrSerValCysCysAlaValCysAsnAla

50 60

ValThrAlaValProProProGlyThrGluMetAlaGlnLeuValCysGlyGlyCysHis

70 80

ThrLeuLeuMetTyrIleArgGlyAlaThrSerValGlnCysSerCysCysHisThrVal

90 100

AsnLeuAlaLeuGluAlaAsnGlnValAlaHisValAsnCysGlyAsnCysMetMetLeu

110 120

LeuMetTyrGlnTyrGlyAlaArgSerValLysCysAlaValCysAsnPheValThrSer

130 140

ValGlyGlySerThrSerThrThrAspSerLysPheAsnAsn

150

SEQ ID NO: 49

CysSerGlyCysArgAsnLeuLeuMetTyrProValGlyAlaThrSerValCysCysAlaValCys

SEQ ID NO: 50

CysGlyGlyCysHisThrLeuLeuMetTyrIleArgGlyAlaThrSerValGlnCysSerCysCys

SEQ ID NO: 51

CysGlyAsnCysMetMetLeuLeuMetTyrGlnTyrGlyAlaArgSerValLysCysAlaValCys

SEQ ID NO: 52

CysXxxXxxCysArgXxxXxxLeuMetTyrXxxXxxGlyAlaSerXxxValXxxCysXxxXxxCys

SEQ ID NO: 53

CysXxxXxxCysXxxXxxLeuLeuMetTyrXxxXxxGlyAlaXxxSerValXxxCysXxxXxxCys

SEQ ID NO:54

10 30 50

GAGGAGGAAGAGGAAGGTCCGCCTCCGGGATGGGAATCTGCAGTTCTTCCTCCTCCAATC

70 90 110

GTCACCATCACCGCCGCCGTAAACCCCAATCCCACCACCGTAGAAATTCCCGAAAAGGCC

14

130 150 170
CAAATGGTATGTGGATCTTGCAGGCGTTTGCTTTCTTATCTAAGAGGATCCAAACATGTT

190 210 230
AAGTGCTCCTCTTGTCAGACTGTTAATCTCGTTCTTGAAGCTAACCAGGTTGGTCAGGTG

250 270 290
AATTGCAACAATTGCAAACTGCTACTGATGTATCCTTATGGAGCTCCAGCTGTTAGATGT

310 330 350
TCCTCCTGCAATTCTGTCACAGATATCAGTGAAAACAACAAACGACCTCCATGGTCTGAG

370 390 410
CAGCAAGGACCACTCAAAAGTTTAAGCAGTCTCAGGAGAGCAGAGAATTAAACTTGAACC

430 450 470
GATTTTTGTCAATTTTGAACCGGTTTGACGACTAAAAACCTTGTAATAATGTCGAAGGAT

490
AGATGAAATAAAATCACACC

SEQ ID NO:55

GluGluGluGluGluGlyProProProGlyTrpGluSerAlaValLeuProProProIle
10 20

ValThrIleThrAlaAlaValAsnProAsnProThrThrValGluIleProGluLysAla
30 40

GlnMetValCysGlySerCysArgArgLeuLeuSerTyrLeuArgGlySerLysHisVal
50 60

LysCysSerSerCysGlnThrValAsnLeuValLeuGluAlaAsnGlnValGlyGlnVal
70 80

AsnCysAsnAsnCysLysLeuLeuLeuMetTyrProTyrGlyAlaProAlaValArgCys
90 100

SerSerCysAsnSerValThrAspIleSerGluAsnAsnLysArgProProTrpSerGlu
110 120

GlnGlnGlyProLeuLysSerLeuSerSerLeuArgArgAlaGluAsn

55

SEQ ID NO:56

CGSCRRLLSYLRGSKHVKCSSC

SEQ ID NO:57

CNNCKLLLMYPYGAPAVRCSSC

SEQ ID NO: 58

```
         10                            30                            50
GGAAGAGATACAACAACAAACGCAGAAGGAAGAACAAAAGCACCGTGAAGAAGAAGAGGA

         70                            90                           110
GGAAGAGGAAGGTCCGCCTCCGGGATGGGAATCTGCAGTTCTTCCTCCTCCAATCGTCAC

        130                           150                           170
CATCACCGCCGCCGTAAACCCCAATCCCACCACCGTAGAAATTCCCGAAAAGGCCCAAAT

        190                           210                           230
GGTATGTGGATCTTGCAGGCGTTTGCTTTCTTATCTAAGAGGATCCAAACATGTTAAGTG

        250                           270                           290
CTCCTCTTGTCAGACTGTTAATCTCGTTCTTGAAGCTAACCAGGTTGGTCAGGTGAATTG

        310                           330                           350
CAACAATTGCAAACTGCTACTGATGTATCCTTATGGAGCTCCAGCTGTTAGATGTTCCTC

        370                           390                           410
CTGCAATTCTGTCACAGATATCAGTGTATGTATTCACAGATGGTTTTGTGCTCCATGTCC

        430                           450                           470
AAATCATACTTGGAAGAGTTGATACATTTTGAGATCCGAGTAAGTAATCATCTGATGAAT

        490                           510                           530
CATTTATAATAAACTGTGTTATATTTCAGGAAAACAACAAACGACCTCCATGGTCTGAGC

        550                           570                           590
AGCAAGGACCACTCAAAAGTTTAAGCAGTCTCAGGAGAGCAGAGAATTAAACTTGAACCG

        610                           630                           650
ATTTTTGTCAATTTTGAACCGGTTTGACGACTAAAAACCTTGTAATAATGTCGAAGGATA

        670                           690
GATGAAATAAAATCACCATTAATAATCTAAAAAAAAAAAAAAAA
```

SEQ ID NO:59

```
         10                            30                            50
CTCTATCCTTACTTCAACGGAGCTTTACCAGACCCAAACTCTCTTAGGCCGCACCGAGAG
```

70 90 110
TTGTTTGTACGTGTGCTTAACGCAGATTACATATGACGCTTCTAACCCACAATTAATTTG

130 150 170
GTTCACTCTTTGCCGCAAACCAAATAGCTCAAAAAAGATTTTAATCCCAATTTCATATCC

190 210 230
TAAATCTGCATCATGGTCGGATAGTGTAGTGGCTGTTGGTCCTAATATCTACGCTATTGG

250 270 290
GGGATTCAGTAATAATAGAACTAAACCTTCGTCTAGCGTCATGGTCATGGATTGTCGTAC

310 330 350
TCACACATGGTGTGAGGCCCCTAGCATGCAGGTTTCCCGTGTGTTCCAATCTACTTGCGT

370 390 410
CCTTGATGGGAAAATATATGTAACAGGAGGCCGCGGAACTCTCGATTCAACGAAATGGAT

430 450 470
GGAGGTTTTTGATACGAAAACCCAAACTTGGGAGTTTTTGCAATTCCCGAGTGAGGAGAA

490 510 530
GATATGCACAGGCTATAAGTGTGAGAGCATAGTGTATGAAGGAACTGTCTATGTAAGGTC

550 570 590
GTATTTTCATAATGTGACTTACAAGCTGCATAAAGGTAGATGGATTCAGCGGCAGACTTT

610 630 650
AGGCGATGAATAATGGATGGCCGTTGCTCATCATTTTTTTGTGTGATAAAGAACGTGTTC

670 690 710
TACTTGTTGCAATAGAAGTGGTAACGGTATGATCGATTGGTATGACTCGGAAAAAGGATC

730 750 770
ATGGACAACTATGAAGGGGTTGGAAAGATTGCCTAAAGTTTATGGTAATGTTAAATTGGC

790 810 830
ATATTATGGTGGAAAAATGGTGGTGGTGTACGTGGAGTGCTAAGGAGTGGGGTAACGTGA

850 870 890
GAAAAATTTGGTGTGCGGAAATTACGATTGAAAAACGCAAGGATGGAGAGATTTGGGGGA

910 930 950
TACTAGAATGGTTTGACGATGTATATAAAGCCAAGGATGAGCTAGAATATTTAGCTGTAG

970 990 1010
TGCATGCTGTTGTTACTACCATCTGATTGATAAGAGAGTCATGTGAACATTGTTCATTGA

1030 1050 1070
TTCACCGATGCAATAACGAATTTATCTACTATCATTTGTTTTGATTTTCTTTCTAAATCT

1090 1110 1130
TTTTTGTTTGTTCTTGTATTGAATTTTACCTTACATTTATTAAGAAAGTCAACTATTTGT

17

```
        1150                 1170                 1190
CAACGTTACTGGAAAGTTAAAAAGGTAAAAGTAATAATAATCTGAGAGTTAACTTTGGAC

        1210                 1230                 1250
ATCTTCGCCGGAGCCGAGACGGAAGGCGTGATGGAAGAGATACAACAACAAACGCAGAAG

        1270                 1290                 1310
GAAGAACAAAAGCACCGTGAAGAAGAAGAGGAGGAAGAGGAAGGTCCGCCTCCGGGATGG

        1330                 1350                 1370
GAATCTGCAGTTCTTCCTCCTCCAATCGTCACCATCACCGCCGCCGTAAACCCCAATCCC

        1390                 1410                 1430
ACCACCGTAGAAATTCCCGGTATTCTTGTAGTCTTGTCTATTTTAGGGTTTATCGATTTG

        1450                 1470                 1490
CTTCCATTTCTTGCTACAGTCTGATCAAATTAGAGATTTTTAGTGGAGTTTGTAGACTTT

        1510                 1530                 1550
TAGAGATAACCCATTTTCGATTCCGAGAATTGATTAGTGTTTTTTTTCTGCAAATCTTCT

        1570                 1590                 1610
TTGTTTTTGGGGTTGTTGCAGAAAAGGCCCAAATGGTATGTGGATCTTGCAGGCGTTTGC

        1630                 1650                 1670
TTTCTTATCTAAGAGGATCCAAACATGTTAAGTGCTCCTCTTGTCAGACTGTTAATCTCG

        1690                 1710                 1730
TTCTTGAAGGTTCGTTCTTCCATGGCTTTTTTATCTCTTATTCATTACTTGAAAAGCTTT

        1750                 1770                 1790
TGTTGATAATCTCAGTCACTTGAAACTCTTAATGGAACAATCTTGGAATGCTCTCTCAGT

        1810                 1830                 1850
CTAGTTTTACTTAGCATGTGTGAATGATATATCTATGTTCTTTTGAGAATCTCAAAATGT

        1870                 1890                 1910
AAGCTTCCTGAGACCAAATGAGTTTAGTTCTTAACTGACACAAGAATGATCTTTGGTTAG

        1930                 1950                 1970
GATTCTTCTCTTAAGCTTTTGTGAGCCTTTTGGTCTCTACTCCATCATAATGTCTCCTTT

        1990                 2010                 2030
GTAGACCATTTATGTGGTCTTTATCCTTTACTCTTACTACTCTTGGGGAAATTGTGTGAT

        2050                 2070                 2090
CTTAAGACCAAGATTGTTCTTCTTAGCTTGTGAATCACTTGGCCTCATTATTGATGAAAT

        2110                 2130                 2150
AGCCTTCTTCTCTTATCGGTTCTGGACTTGTCGTTCTTTGTTTGCAGCTAACCAGGTTGG

        2170                 2190                 2210
TCAGGTGAATTGCAACAATTGCAAACTGCTACTGATGTATCCTTATGGAGCTCCAGCTGT

        2230                 2250                 2270
```

TAGATGTTCCTCCTGCAATTCTGTCACAGATATCAGTGTATGTATTCACAGATGGTTTTG

2290 2310 2330
TGCTCCATGTCCAAATCATACTTGGAAGAGTTGATACATTTTGAGATCCGAGTAAGTAAT

2350 2370 2390
CATCTGATGAATCATTTATAATAAACTGTGTTATATTTCAGGAAAACAACAAACGACCTC

2410 2430 2450
CATGGTCTGAGCAGCAAGGACCACTCAAAAGTTTAAGCAGTCTCAGGAGAGCAGAGAATT

2470 2490 2510
AAACTTGAACCGATTTTTGTCAATTTTGAACCGGTTTGACGACTAAAAACCTTGTAATAA

2530 2550 2570
TGTCGAAGGATAGATGAAATAAAATCACCATTAATAATCTCATTGAATTCCCATTCTTTC

2590 2610 2630
AGATATTACTTGCTCATCATCCTTTACTGTTTTAAGCTTTAGTGGTTAAAAAGAATGTGT

2650 2670 2690
ATATATCCATACAAAAGTTGATATATGTACTGGACCAATATAAACAAACACAGCTCACAG

2710 2730 2750
TCTCACACAATACATAAAAACAAAATTCATATTTCACAGGTGAGAAAAACTAACTAGTAG

2770 2790 2810
TCTACTTGGCCGAATTTGTCAATGAATTTCAATAATTAGGTCGTATAAATAGCAAACAAA

2830 2850 2870
ACATGGACTCTTACCCAACCAAATATGCATAAATAATTTACATTACAGTTTCATATAAAA

2890 2910 2930
TACAAACTAATGGTGGGTCCTCGAGAGAGCTAACAAGAGCTGTGTGTGGGTGAAGAACCA

2950 2970 2990
ACTTGTCAACGAAACCAATTTAATGGAAATCAACCCTAAATTTAATGAAACCTTGGACGA

3010 3030 3050
AACTTACATTTTGTTAACCAGTTTATCCTTTTAAATCAAACCTGCATAGAATTTTGATTT

SEQ ID NO:60

MetGluGluIleGlnGlnGlnThrGlnLysGluGluGlnLysHisArgGluGluGluGlu
10 20

GluGluGluGluGlyProProProGlyTrpGluSerAlaValLeuProProProIleVal
30 40

ThrIleThrAlaAlaValAsnProAsnProThrThrValGluIleProGluLysAlaGln
50 60

MetValCysGlySerCysArgArgLeuLeuSerTyrLeuArgGlySerLysHisValLys
70 80

CysSerSerCysGlnThrValAsnLeuValLeuGluAlaAsnGlnValGlyGlnValAsn
90 100

CysAsnAsnCysLysLeuLeuLeuMetTyrProTyrGlyAlaProAlaValArgCysSer
110 120

SerCysAsnSerValThrAspIleSerGluAsnAsnLysArgProProTrpSerGluGln
130 140

GlnGlyProLeuLysSerLeuSerSerLeuArgArgAlaGluAsn
150

SEQ ID NO:61

MetGluGluIleGlnGlnGlnThrGlnLysGluGluGlnLysHisArgGluGluGluGlu
10 20

GluGluGluGluGlyProProProGlyTrpGluSerAlaValLeuProProProIleVal
30 40

ThrIleThrAlaAlaValAsnProAsnProThrThrValGluIleProGluLysAlaGln
50 60

MetValCysGlySerCysArgArgLeuLeuSerTyrLeuArgGlySerLysHisValLys
70 80

CysSerSerCysGlnThrValAsnLeuValLeuGluAlaAsnGlnValGlyGlnValAsn
90 100

CysAsnAsnCysLysLeuLeuLeuMetTyrProTyrGlyAlaProAlaValArgCysSer
110 120

SerCysAsnSerValThrAspIleSerValCysIleHisArgTrpPheCysAlaProCys
130 140

ProAsnHisThrTrpLysSer

SEQ ID NO:62

CysXxxXxxCysXxxXxxLeuLeuXxxTyrXxxXxxGlyXxxXxxXxxValXxxCysSerSerCys

SEQ ID NO: 63

LeuValCysHisGlyCysArgAsnLeuLeuMetTyrProArgGlyAlaSerAsnValArgCysAlaLeuCysA
snThrIleAsnMetVal
IleIleCysGlyGlyCysArgThrMetLeuMetTyrThrArgGlyAlaSerSerValArgCysSerCysCysG
lnThrThrAsnLeuVal

20

IleAsnCysGlyHisCysArgThrThrLeuMetTyrProTyrGlyAlaSerSerValLysCysAlaValCysG
lnPheValThrAsnVal

SEQ ID NO: 64

LeuValCysSerGlyCysArgAsnLeuLeuMetTyrProValGlyAlaThrSerValCysCysAlaValCysA
snAlaValThrAlaVal
LeuValCysGlyGlyCysHisThrLeuLeuMetTyrIleArgGlyAlaThrSerValGlnCysSerCysCysH
isThrValAsnLeuAla
ValAsnCysGlyAsnCysMetMetLeuLeuMetTyrGlnTyrGlyAlaArgSerValLysCysAlaValCysA
snPheValThrSerVal

SEQ ID NO: 65

MetValCysGlySerCysArgArgLeuLeuSerTyrLeuArgGlySerLysHisValLysCysSerSerCysG
lnThrValAsnLeuVal
ValAsnCysAsnAsnCysLysLeuLeuLeuMetTyrProTyrGlyAlaProAlaValArgCysSerSerCysA
snSerValThrAspIle

SEQ ID NO: 66

Nucleic acid sequence of C

```
       10                   30                   50
AGCAACAACAACAACAACCAGCAACCACCACCAACCTCCGTCTATCCACCTGGCTCCGCC

       70                   90                  110
GTCACAACCGTAATCCCTCCTCCACCATCTGGATCTGCATCAATAGTCACCGGAGGAGGA

      130                  150                  170
GCGACATACCACCACCTCCTCCAGCAACAACAGCAACAGCTTCAAATGTTCTGGACATAC

      190                  210                  230
CAGAGACAAGAGATCGAACAGGTAAACGATTTCAAAAACCATCAGCTCCCTCTAGCTCGT

      250                  270                  290
ATCAAAAAAATCATGAAAGCTGATGAAGATGTGCGTATGATCTCCGCCGAAGCACCGATT

      310                  330                  350
CTCTTCGCGAAAGCTTGTGAGCTTTTCATTCTCGAACTTACGATTAGATCTTGGCTTCAC

      370                  390                  410
GCTGAAGAGAACAAACGTCGTACGCTTCAGAAAAACGATATCGCTGCTGCGATTACTAGA

      430                  450                  470
ACCGATATCTTCGATTTCCTTGTTGATATTGTTCCTAGGGAAGAGATCAAGGAAGAGGAA

      490                  510                  530
GATGCAGCATCGGCTCTTGGTGGAGGAGGTATGGTTGCTCCCGCCGCGAGCGGTGTTCCT
```

21

```
        550                 570                 590
TATTATTATCCACCGATGGGACAACCGGCGGTTCCTGGAGGGATGATGATTGGAAGACCG

        610                 630                 650
GCGATGGATCCTAGCGGTGTTTATGCTCAGCCTCCTTCTCAGGCATGGCAAAGCGTTTGG

        670                 690                 710
CAGAATTCAGCTGGTGGTGGTGATGATGTGTCTTATGGAAGTGGAGGAAGTAGCGGCCAT

        730                 750                 770
GGTAATCTCGATAGCCAAGGTTGAGCTATGGAACCAGAAGCTTAGAGATTTAATCATCAT

        790                 810                 830
TTCGACCCTGCAAGTGTCTGATTCTTATATGTCTATGATTCGAATGACTTA
```

SEQ ID NO: 67

Amino acid sequence of C

```
SerAsnAsnAsnAsnAsnGlnGlnProProProThrSerValTyrProProGlySerAla
                              10                            20

ValThrThrValIleProProProProSerGlySerAlaSerIleValThrGlyGlyGly
                              30                            40

AlaThrTyrHisHisLeuLeuGlnGlnGlnGlnGlnGlnLeuGlnMetPheTrpThrTyr
                              50                            60

GlnArgGlnGluIleGluGlnValAsnAspPheLysAsnHisGlnLeuProLeuAlaArg
                              70                            80

IleLysLysIleMetLysAlaAspGluAspValArgMetIleSerAlaGluAlaProIle
                              90                           100

LeuPheAlaLysAlaCysGluLeuPheIleLeuGluLeuThrIleArgSerTrpLeuHis
                             110                           120

AlaGluGluAsnLysArgArgThrLeuGlnLysAsnAspIleAlaAlaAlaIleThrArg
                             130                           140

ThrAspIlePheAspPheLeuValAspIleValProArgGluGluIleLysGluGluGlu
                             150                           160

AspAlaAlaSerAlaLeuGlyGlyGlyGlyMetValAlaProAlaAlaSerGlyValPro
                             170                           180

TyrTyrTyrProProMetGlyGlnProAlaValProGlyGlyMetMetIleGlyArgPro
                             190                           200

AlaMetAspProSerGlyValTyrAlaGlnProProSerGlnAlaTrpGlnSerValTrp
                             210                           220
```

```
GlnAsnSerAlaGlyGlyGlyAspAspValSerTyrGlySerGlyGlySerSerGlyHis
                                    230                                     240

GlyAsnLeuAspSerGlnGly
```

SEQ ID NO: 68

Nucleic acid sequence of CC

```
        10                          30                          50
AGTATGGATGAGCTTTCAGAAGCTTCTCAGATACTCACATGTTGCTCTGACATGGTGTAC

        70                          90                          110
TGCACGGTTTGCGCATGTATGCAGACACAACACAAGATGGAAATGGACAAGAGGGACGGT

        130                         150                         170
AAGTTCGGGCCACAGCCAATGGCAGTGCCTCCGGCTCAGCAAATGTCACGGTTTGATCAA

        190                         210                         230
GCCACCCCACCCGCAGTCGGTTATCCTCCACAACAAGGTTATCCACCTTCTGGTTATCCT

        250                         270                         290
CAACACCCTCCACAAGGTTATCCACCTTCTGGCTATCCTCAAAACCCTCCTCCCTCAGCT

        310                         330                         350
TATTCTCAATACCCTCCTGGGGCTTATCCTCCTCCTCCCGCTTACCCAAAGTGATCACTC

        370                         390                         410
TTTGCCTGTTTTCTCTCCCGATTGGAAAATTTTATTTCATCTTTTTTTAATGCTGTCTTG

        430                         450                         470
TTACGGGTCAAGAATTGAACGTTCGCTGATTGTTTTGAGGTCGTTGTTTGTATGAGATTT

        490                         510                         530
TGACCTCGCATGTTGTTGTTGTTTTCTGAAACGTCCCTCTTGGACTAAGAGATTTCATGA

        550
CTTAAAAAAAAAAAAAAAAAA
```

SEQ ID NO: 69

Amino acid sequence of CC

```
SerMetAspGluLeuSerGluAlaSerGlnIleLeuThrCysCysSerAspMetValTyr
                                    10                                      20

CysThrValCysAlaCysMetGlnThrGlnHisLysMetGluMetAspLysArgAspGly
```

23

30 40

LysPheGlyProGlnProMetAlaValProProAlaGlnGlnMetSerArgPheAspGln

50 60

AlaThrProProAlaValGlyTyrProProGlnGlnGlyTyrProProSerGlyTyrPro

70 80

GlnHisProProGlnGlyTyrProProSerGlyTyrProGlnAsnProProProSerAla

90 100

TyrSerGlnTyrProProGlyAlaTyrProProProProAlaTyrProLys

110

SEQ ID NO: 70

Nucleic acid sequence of FF

10 30 50

AGGTTTCCGACGTTGATGACCCAATTTCCGTCGTCGACGAAGACGATTCCGGCATCGTAT

70 90 110

TTGCTTCCGTTACAATGGCCTCAGCCGCAGAACGAGGAGATTCTTCTCGCCATGGAAGAA

130 150 170

GCTGAGTTCGAAGAAAAGTGCAACGAGATCAGAAAGATGAGTCCTGCTTTACCGGTAATT

190 210 230

GGAAAACCAGTCGTCAACAACGAACAAGAAGAGGATGATAATGAATCAGAGGATGATGAT

250 270 290

GCAGATAATGCAGAGGAATCAGATGGTGAAGAGTTTGAGCAAGAAACCGGATAAATAATC

310 330 350

TTGAGGCCGAAAATACACAAGGGTTATTGATGGCATTGGCTTGAAACTTGAGGACCCTTA

370 390 410

TCTAAATCTTCTTGTGATAAAACGACTGTGATTCTGACTTTGTAAACCANGTTTTTTTCT

430 450 470

TTTCTTAGGAACGACTGAAATGTTCACTTTTGGCCCTAAGGTTAGTCAGTGGATTATTCG

490 510 530

TAGTTAATTGTCTCAATCTCATGGTGTTAATTGTGTTAGTGTATTGACATTGAATTTTAT

550 570

GGTTTATAGATTGTAGTGATTTGATGAAAAAAAAAAAAAAAAAAAAA

24

SEQ ID NO: 71

Amino acid sequence of FF

```
ArgPheProThrLeuMetThrGlnPheProSerSerThrLysThrIleProAlaSerTyr
                    10                                      20

LeuLeuProLeuGlnTrpProGlnProGlnAsnGluGluIleLeuLeuAlaMetGluGlu
                    30                                      40

AlaGluPheGluGluLysCysAsnGluIleArgLysMetSerProAlaLeuProValIle
                    50                                      60

GlyLysProValValAsnAsnGluGlnGluGluAspAspAsnGluSerGluAspAspAsp
                    70                                      80

AlaAspAsnAlaGluGluSerAspGlyGluGluPheGluGlnGluThrGly
                    90
```

SEQ ID NO: 72

Nucleic acid sequence of GG

```
        10                  30                  50
AGGGAAACAATGAGCCAGTACAATCAACCTCCCGTTGGTGTTCCTCCTCCTCAAGGTTAT

        70                  90                 110
CCACCGGAGGGATATCCAAAAGATGCTTATCCACCACAAGGATATCCTCCTCAGGGATAT

       130                 150                 170
CCTCAGCAAGGCTATCCACCTCAGGGATATCCTCAACAAGGTTATCCTCAGCAAGGATAT

       190                 210                 230
CCTCCACCGTACGCGCCTCAATATCCTCCACCACCGCAGCATCAGCAACAACAGAGCAGT

       250                 270                 290
CCTGGCTTTCTAGAAGGATGTCTTGCTGCTCTGTGTTGTTGCTGTCTCTTGGATGCTTGC

       310                 330                 350
TTCTGATTGGAGTCTCTCTCTCTCTGCATAAAGCTTCGGGATTTATTTGTAAGAGGGTTT

       370                 390                 410
TGGTTAAACAAAAACCTTAATTGATTTGTGGGGCATTAAAAATGAATCTCTCGATGATTC

       430                 450                 470
TCTTTCGTTTTATGTGTAATGTTCTTCGGTTCATAACATTTTAACTATTGTCTATCGACG

       490                 510                 530
TTCTGCCTTAGTTTGTATTTGATTATGGGAATGTAAATTGGTTGGGAGACACTATTCTAT
```

550 570
GCCATAGTTTATTGCTTGGATCTTCAAAAAAAAAAAAAAAAA

SEQ ID NO: 73

Amino acid sequence of GG

ArgGluThrMetSerGlnTyrAsnGlnProProValGlyValProProProGlnGlyTyr
10 20

ProProGluGlyTyrProLysAspAlaTyrProProGlnGlyTyrProProGlnGlyTyr
30 40

ProGlnGlnGlyTyrProProGlnGlyTyrProGlnGlnGlyTyrProGlnGlnGlyTyr
50 60

ProProProTyrAlaProGlnTyrProProProProGlnHisGlnGlnGlnGlnSerSer
70 80

ProGlyPheLeuGluGlyCysLeuAlaAlaLeuCysCysCysCysLeuLeuAspAlaCys
90 100

Phe

SEQ ID NO: 74

Nucleic acid sequence of HH

10 30 50
AGTGATGTTCTTCCTAAGTCCGTTGACTGGAGAAACGAAGGCGCAGTGACTGAAGTCAAA

70 90 110
GATCAAGGCCTTTGCAGGAGTTGTTGGGCTTTCTCCACTGTGGGAGCAGTGGAAGGCTTA

130 150 170
AACAAGATTGTGACTGGAGAGCTAGTAACTTTGTCTGAGCAAGATTTGATCAATTGTAAC

190 210 230
AAAGAAAACAATGGTTGCGGAGGAGGCAAAGTCGAGACAGCCTATGAGTTCATCATGAAC

250 270 290
AATGGTGGTCTTGGTACCGACAACGATTATCCTTACAAAGCTCTCAATGGAGTCTGCGAA

310 330 350
GGCCGCCTCAAGGAAGACAACAAGAATGTTATGATTGATGGGTATGAGAATTTGCCTGCA

370 390 410
AACGATGAAGCCGCTCTCATGAAAGCGGTTGCTCACCAGCCTGTGACTGCCGTTGTCGAT

```
      430                           450                           470
TCCAGCAGCCGAGAGTTTCAGCTTTATGAATCGGGAGTGTTTGACGGAACTTGCGGAACA

      490                           510                           530
AACCTAAACCATGGTGTTGTTGTGGTCGGGTATGGAACCGAGAATGGTCGTGACTACTGG

      550                           570                           590
ATTGTGAAAAACTCGAGGGGCGACACATGGGGGGAGGCTGGCTACATGAAGATGGCTCGC

      610                           630                           650
AACATTGCCAATCCAAGAGGCATATGTGGCATCGCAATGCGAGCTTCATACCCTCTCAAG

      670                           690                           710
AACTCGTTTTCTACGGATAAAGTTTCGGTTGCCTAATAATATGAACTAAATGTATGCCAT

      730                           750                           770
GGAACGGATCGGTTAAGCCATTATCGTTATTCGACTTTGAAGGAAACTAAAAAATAATGT

      790                           810                           830
GGTCGATTGGTTTGGTTTTGTTATATATTATGCATTTGTATGGGGGTCAGTCAATGTTTG

      850                           870                           890
AACTTTGTATAATATTTCTTTGGGTCTAGTGATAAATATTTTCCCTTTTGCGAAAAAAAA

      910
AAAAAAAAAA
```

SEQ ID NO: 75

Amino acid sequence of HH

```
SerAspValLeuProLysSerValAspTrpArgAsnGluGlyAlaValThrGluValLys
                            10                                20

AspGlnGlyLeuCysArgSerCysTrpAlaPheSerThrValGlyAlaValGluGlyLeu
                            30                                40

AsnLysIleValThrGlyGluLeuValThrLeuSerGluGlnAspLeuIleAsnCysAsn
                            50                                60

LysGluAsnAsnGlyCysGlyGlyGlyLysValGluThrAlaTyrGluPheIleMetAsn
                            70                                80

AsnGlyGlyLeuGlyThrAspAsnAspTyrProTyrLysAlaLeuAsnGlyValCysGlu
                            90                               100

GlyArgLeuLysGluAspAsnLysAsnValMetIleAspGlyTyrGluAsnLeuProAla
                           110                               120

AsnAspGluAlaAlaLeuMetLysAlaValAlaHisGlnProValThrAlaValValAsp
                           130                               140
```

27

SerSerSerArgGluPheGlnLeuTyrGluSerGlyValPheAspGlyThrCysGlyThr
150 160

AsnLeuAsnHisGlyValValValValGlyTyrGlyThrGluAsnGlyArgAspTyrTrp
170 180

IleValLysAsnSerArgGlyAspThrTrpGlyGluAlaGlyTyrMetLysMetAlaArg
190 200

AsnIleAlaAsnProArgGlyIleCysGlyIleAlaMetArgAlaSerTyrProLeuLys
210 220

AsnSerPheSerThrAspLysValSerValAla
230

SEQ ID NO: 76

Nucleic acid sequence of I

10 30 50
AGCGAAATGCCAGTTTCAGCTCCATCTCCGCCTCGTCTTCATTCTCCGTTCATTCACTGT

70 90 110
CCCATCAATTTCACTCCTTCTTCTTTCTCGGCGAGGAATCTCCGGTCGCCGTCAACATCT

130 150 170
TATCCCCGAATCAAAGCTGAACTCGATCCCAACACGGTAGTCGCGATATCTGTAGGCGTA

190 210 230
GCAAGCGTCGCATTAGGAATCGGAATCCCTGTGTTCTACGAGACTCAAATCGACAATGCG

250 270 290
GCTAAGCGAGAGAATACTCAACCTTGTTTTCCCTGTAATGGCACCGGAGCTCAGAAATGC

310 330 350
AGATTGTGTGTGGGAAGTGGTAATGTGACCGTAGAGCTTGGTGGAGGAGAGAAAGAAGTC

370 390 410
TCAAACTGTATCAACTGTGATGGTGCTGGTTCCTTAACTTGCACTACTTGTCAAGGCTCT

430 450 470
GGTGTTCAACCTCGATACCTTGATCGAAGGGAGTTCAAGGACGATGACTAAATACCTTGC

490 510 530
TCTAAGGAACATTTCTTTTCTTCTCCCTTCTCACATTTCTTCATTGTACAATGCTGTTTT

550 570 590
GTTCACCAAACATGTTGAGAGAACATCATGACATGGATATTGTAATTGTGAAAGAAAACC

610 630 650

28

ACCAGAGTTCAATCAAATGTTTCTTCTTGTACTTAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 77

Amino acid sequence of I

SerGluMetProValSerAlaProSerProProArgLeuHisSerProPheIleHisCys
10 20

ProIleAsnPheThrProSerSerPheSerAlaArgAsnLeuArgSerProSerThrSer
30 40

TyrProArgIleLysAlaGluLeuAspProAsnThrValValAlaIleSerValGlyVal
50 60

AlaSerValAlaLeuGlyIleGlyIleProValPheTyrGluThrGlnIleAspAsnAla
70 80

AlaLysArgGluAsnThrGlnProCysPheProCysAsnGlyThrGlyAlaGlnLysCys
90 100

ArgLeuCysValGlySerGlyAsnValThrValGluLeuGlyGlyGlyGluLysGluVal
110 120

SerAsnCysIleAsnCysAspGlyAlaGlySerLeuThrCysThrThrCysGlnGlySer
130 140

GlyValGlnProArgTyrLeuAspArgArgGluPheLysAspAspAsp
150

SEQ ID NO: 78

Nucleic acid sequence of II

10 30 50
AGAGAAAACATGGGAGGTGACAATGATAATGACAAAGACAAAGGGTTTCATGGGTATCCT

70 90 110
CCCGCTGGATACCCACCCCCTGGGGCTTATCCACCCGCTGGATACCCACAACAAGGTTAC

130 150 170
CCTCCACCACCCGGTGCTTACCCGCCTGCAGGTTATCCTCCGGGTGCCTACCCACCTGCT

190 210 230
CCTGGTGGTTATCCTCCCGCCCCTGGTTATGGTGGTTATCCTCCAGCTCCTGGTTATGGA

250 270 290

29

GGTTATCCTCCTGCACCTGGTCATGGTGGTTACCCTCCTGCTGGCTATCCTGCTCATCAC

310 330 350
TCAGGACACGCAGGAGGAATTGGGGGTATGATTGCAGGTGCTGCAGCTGCCTATGGAGCT

370 390 410
CACCACGTATCTCATAGCTCTCACTGTCCTTACGGACATGCTGCATATGGTCACGGTTTT

430 450 470
GGCCATGGTCATGGCTATGGCTATGGTCATGGTCATGGTAAGTTCAAGCATGGGAAGCAC

490 510 530
GGGAAGTTCAAGCATGGGAAGCATGGAATGTTTGGAGGAGGCAAGTTCAAGAAGTGGAAG

550 570 590
TGATCTAGCTATTACCTTGTGTGAATTTGTCTGGACTGACCAATGTTTCAAATAAGCCCT

610 630 650
AAACATTATATAAGTTGACTTTCGTCGGTTAGATTGCTGGTTCGAGTTGGAATAATTGAA

670 690 710
ACTTAATTAGTATCAAATCTTATTGTGTACTTTAAAGCTATCGTTGGCTTTATAATGACA

730 750 770
GATTCTGGTTTCGGTGTGTTGTTTTAAGATTTTTGTATATACTGTTTTTTACATTGCTTA

790 810
AGCTTATAGAAGTCATGATTATGATTAAAAAAAAAAAAAAAAAA

SEQ ID NO: 79

Amino acid sequence of II

ArgGluAsnMetGlyGlyAspAsnAspAsnAspLysAspLysGlyPheHisGlyTyrPro
10 20

ProAlaGlyTyrProProProGlyAlaTyrProProAlaGlyTyrProGlnGlnGlyTyr
30 40

ProProProProGlyAlaTyrProProAlaGlyTyrProProGlyAlaTyrProProAla
50 60

ProGlyGlyTyrProProAlaProGlyTyrGlyGlyTyrProProAlaProGlyTyrGly
70 80

GlyTyrProProAlaProGlyHisGlyGlyTyrProProAlaGlyTyrProAlaHisHis
90 100

SerGlyHisAlaGlyGlyIleGlyGlyMetIleAlaGlyAlaAlaAlaAlaTyrGlyAla
110 120

HisHisValSerHisSerSerHisCysProTyrGlyHisAlaAlaTyrGlyHisGlyPhe

```
                        130                                 140
GlyHisGlyHisGlyTyrGlyTyrGlyHisGlyHisGlyLysPheLysHisGlyLysHis
                        150                                 160
GlyLysPheLysHisGlyLysHisGlyMetPheGlyGlyGlyLysPheLysLysTrpLys
                        170                                 180
```

SEQ ID NO: 80

Nucleic acid sequence of K

```
         10                    30                    50
AGTGTCACTACTCCATCCGAGGAGGATTCAAACAACGGTTTACCGGTTCAGCAACCCGGT

         70                    90                   110
ACACCGAACCAGCGAACCAGAGTTCCCGTGAGTCAATTCGCGCCGCCGAATTATCAGCAA

        130                   150                   170
GCTAATGTTAACCTATCTGTTGGGAGGCCATGGAGCACTGGTTTGTTTGATTGTCAAGCA

        190                   210                   230
GACCAAGCCAATGCCGTTTTGACCACAATTGTACCTTGTGTAACATTTGGACAAATAGCA

        250                   270                   290
GAAGTGATGGATGAAGGAGAGATGACTTGTCCTCTTGGAACTTTCATGTACTTATTGATG

        310                   330                   350
ATGCCGGCTTTATGCTCTCACTGGGTGATGGGATCAAAGTATAGAGAAAAAATGAGGAGA

        370                   390                   410
AAATTTAATCTTGTGGAAGCTCCATATTCAGATTGTGCCAGTCATGTCCTATGCCCTTGT

        430                   450                   470
TGCTCTCTTTGTCAAGAATACAGAGAGCTCAAGATTAGGAATCTTGATCCTTCTCTAGGT

        490                   510                   530
TGGAATGGGATACTTGCTCAAGGACAAGGACAATATGAGAGAGAAGCACCAAGTTTTGCT

        550                   570                   590
CCTACAAATCAATATATGTCTAAGTAAACATTTGATTTTAGTTGACTTCCATATTTATTA

        610                   630                   650
AAACATTATTTGTGGACCATTGTACAATGAAAGTGTGCTATATTAAAATTTGCAATGCAA

        670                   690
GTGTGAGATTGATAAAAAAAAAAAAAAAAAAA
```

SEQ ID NO: 81

31

Amino acid sequence of K

SerValThrThrProSerGluGluAspSerAsnAsnGlyLeuProValGlnGlnProGly
10 20

ThrProAsnGlnArgThrArgValProValSerGlnPheAlaProProAsnTyrGlnGln
30 40

AlaAsnValAsnLeuSerValGlyArgProTrpSerThrGlyLeuPheAspCysGlnAla
50 60

AspGlnAlaAsnAlaValLeuThrThrIleValProCysValThrPheGlyGlnIleAla
70 80

GluValMetAspGluGlyGluMetThrCysProLeuGlyThrPheMetTyrLeuLeuMet
90 100

MetProAlaLeuCysSerHisTrpValMetGlySerLysTyrArgGluLysMetArgArg
110 120

LysPheAsnLeuValGluAlaProTyrSerAspCysAlaSerHisValLeuCysProCys
130 140

CysSerLeuCysGlnGluTyrArgGluLeuLysIleArgAsnLeuAspProSerLeuGly
150 160

TrpAsnGlyIleLeuAlaGlnGlyGlnGlyGlnTyrGluArgGluAlaProSerPheAla
170 180

ProThrAsnGlnTyrMetSerLys

SEQ ID NO: 82

Nucleic acid sequence of M

10 30 50
AGAAAATACGAAAAGGTCTCCCTCCCAGCACCTTACGTGGCTGGACACTCGAGCCATCAC

70 90 110
GAAGACGACGGTCAATACTATCCCGGCAAATACGAAAAAGCCTCCCTCCCAGCACCTTAC

130 150 170
GTGGCCGGATATCCGAGCCATCATGAAGACGATGGTCAATACTATCCTGGCAAATACGAA

190 210 230
AAGGTCTCCCTCCCAGCACCTTACGTGGTCGGACACCCGAGCCACTCCGAAGATGATGGC

250 270 290
CAATACTATCCCGGCAAATACGAAAAGGCCTCCGTCCCATCAGCTTACGTGGCCGAACAC

310 330 350
TCGAGCCACTCCGAAGATGATGGCCAATACTATCCTGGCAAATACGAAAAGCCCGAACAC

370 390 410
CATTACTGAAAACTCTCACACAACAATGATTCTCATCCTTCCGTAGTCTTTTAATTCGAC

430 450 470
TTTTAACAATAAAAACGTGATCTTAATTTTTCATCAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 83

Amino acid sequence of M

ArgLysTyrGluLysValSerLeuProAlaProTyrValAlaGlyHisSerSerHisHis
10 20

GluAspAspGlyGlnTyrTyrProGlyLysTyrGluLysAlaSerLeuProAlaProTyr
30 40

ValAlaGlyTyrProSerHisHisGluAspAspGlyGlnTyrTyrProGlyLysTyrGlu
50 60

LysValSerLeuProAlaProTyrValValGlyHisProSerHisSerGluAspAspGly
70 80

GlnTyrTyrProGlyLysTyrGluLysAlaSerValProSerAlaTyrValAlaGluHis
90 100

SerSerHisSerGluAspAspGlyGlnTyrTyrProGlyLysTyrGluLysProGluHis
110 120

HisTyr

SEQ ID NO: 84

Nucleic acid sequence of OO

10 30 50
AGCCGATCTCAGATTCTTCCATCTTCCAGGAGGAATTTCAGTGTGGCGACCACACAGCTT

70 90 110
GGCATTCCAACAGACGATCTAGTCGGCAATCACACCGCCAAATGGATGCAGGATAGAAGC

130 150 170
AAGAAATCACCTATGGAACTGATTAGTGAGGTTCCACCTATCAAAGTTGATGGAAGGATT

190 210 230
GTTGCTTGTGAAGGAGACACCAATCCGGCCCTAGGTCATCCAATCGAGTTCATATGCCTC

33

```
       250                 270                 290
GACCTAAATGAGCCTGCGATCTGCAAGTACTGCGGCCTTCGTTATGTTCAAGATCATCAC

       310                 330                 350
CATTGAGGCAAATTCTGAAAGTGAACTGCTGGTCTCTCTCCCCTTTTTATTGCATTTTTA

       370                 390                 410
AGTTTGTGTATTGTTTTTTTCTGGTGTGCCTACTACATCTTCAGCTATATTATCTAATAA

       430                 450                 470
AGGATTCGATCAAAGTCGGGTAAGTTTGATTTTTGTTTGATCTCACTTCAGCACTTGTCA

       490                 510                 530
TGTTGTAACATTCAATCTCTGATATCACTGTCTTTTACATGCCAAAAAAAAAAAAAAAAA

       550
AAAAAAAAAAAAAAAA
```

SEQ ID NO: 85

Amino acid sequence of OO

```
SerArgSerGlnIleLeuProSerSerArgArgAsnPheSerValAlaThrThrGlnLeu
                              10                            20

GlyIleProThrAspAspLeuValGlyAsnHisThrAlaLysTrpMetGlnAspArgSer
                              30                            40

LysLysSerProMetGluLeuIleSerGluValProProIleLysValAspGlyArgIle
                              50                            60

ValAlaCysGluGlyAspThrAsnProAlaLeuGlyHisProIleGluPheIleCysLeu
                              70                            80

AspLeuAsnGluProAlaIleCysLysTyrCysGlyLeuArgTyrValGlnAspHisHis
                              90                           100

HisEndGlyLysPhe
```

SEQ ID NO: 86

Nucleic acid sequence of P

```
       10                  30                  50
AGAACAGCTCGAGTTCCTTATGGGCCTAGACTCTCTGGTGGTGGTTACAACCGATCTGGA

       70                  90                 110
AACAGGGTTCCGCGTAACAAACCAAGCTTCCCCAATAGCACCGAGTCCAATGGTGAGGCT
```

34

```
      130                 150                 170
AATCAATTCAATGGCCCAAGAATAATGAACCCCCATGCTGCTGAGTTCATACCGAGTCAA

      190                 210                 230
CCTTGGGTTTCTAATGGGTATCCAGTGTCACCAAATGGCTATTTAGCATCCCCAAATGGT

      250                 270                 290
GCAGAAATAACACAGAATGGGTACCCTTTGTCACCAGTAGCAGGTGGATATCCGTGTAAC

      310                 330                 350
ATGTCCGTTACACAGCCTCAGGATGGACTTGTTTCAGAGGAATTACCTGGTGCTGGAAGC

      370                 390                 410
TCTGAGGAGAAGAGCGGAAGCGAAGAAGAAAGCAACAACGACAAAAATGCTGGAGAGGAT

      430                 450                 470
GACGAAGCCGTTGGACAAGAAACTACAGATACACCTGAAAATGGACATTCGACAGTAGGT

      490                 510                 530
GAAGTGGAAACCACATCACATGAGACTTGTGATGAGAAAAATGGAGAACGACAAGGAGGC

      550                 570                 590
AAGTGCTGGGGAGATTACAGCGATAATGAAATCGAGCAAATTGAAGTTACAAGTTGAAGA

      610                 630                 650
CGCAACTGTCTGTTACTGAAGTATTAACATTGAGGCTAAAGGAATGCGGAGACATTTTGG

      670                 690                 710
CTCCATTGATGAGGTTAAAGGTAAACAATCATCATAGTCGAGAAAAGCATTTTTACATGT

      730                 750                 770
GAATGTTTTGTGTTGTAGCGCAGGACCAAGGCTCGTCACTCCTGCTTTAACAACTTTTCT

      790                 810                 830
CCTGCTTTCAGTTTTTGGTTTCATAGCTGAAAACTAGATATATTCAACTCCTTAATAAAA

      850                 870
GATTTGTCCCTTTGTTTAAAAAAAAAAAAAAAAAAAAAA
```

SEQ ID NO: 87

Amino acid sequence of P

```
ArgThrAlaArgValProTyrGlyProArgLeuSerGlyGlyGlyTyrAsnArgSerGly
                             10                            20

AsnArgValProArgAsnLysProSerPheProAsnSerThrGluSerAsnGlyGluAla
                             30                            40

AsnGlnPheAsnGlyProArgIleMetAsnProHisAlaAlaGluPheIleProSerGln
                             50                            60

ProTrpValSerAsnGlyTyrProValSerProAsnGlyTyrLeuAlaSerProAsnGly
```

70 80

AlaGluIleThrGlnAsnGlyTyrProLeuSerProValAlaGlyGlyTyrProCysAsn
90 100

MetSerValThrGlnProGlnAspGlyLeuValSerGluGluLeuProGlyAlaGlySer
110 120

SerGluGluLysSerGlySerGluGluGluSerAsnAsnAspLysAsnAlaGlyGluAsp
130 140

AspGluAlaValGlyGlnGluThrThrAspThrProGluAsnGlyHisSerThrValGly
150 160

GluValGluThrThrSerHisGluThrCysAspGluLysAsnGlyGluArgGlnGlyGly
170 180

LysCysTrpGlyAspTyrSerAspAsnGluIleGluGlnIleGluValThrSer
190

SEQ ID NO: 88

Nucleic acid sequence of T

10 30 50
AGAGACCATCCAGCTTACCATCAGATCCACCAGCAACAACAACAACAGCTCACTCAACAG

70 90 110
CTTCAATCTTTCTGGGAGACTCAATTCAAAGAGATTGAGAAAACCACTGATTTCAAGAAC

130 150 170
CATAGCCTTCCATTGGCAAGAATCAAGAAAATCATGAAAGCTGATGAAGATGTGCGTATG

190 210 230
ATCTCGGCCGAGGCGCCTGTTGTGTTCGCCAGGGCCTGCGAGATGTTTATTCTGGAGCTT

250 270 290
ACGTTAAGGTCTTGGAACCATACTGAGGAGAACAAGAGAAGGACGTTGCAGAAGAATGAT

310 330 350
ATCGCGGCTGCGGTGACTAGAACTGATATTTTTGATTTTCTTGTGGATATTGTTCCTCGG

370 390 410
GAGGATCTTCGTGATGAAGTCTTGGGTGGTGTTGGTGCTGAAGCTGCTACAGCTGCGGGT

430 450 470
TATCCGTATGGATACTTGCCTCCTGGAACAGCTCCAATTGGGAACCCGGGAATGGTTATG

490 510 530
GGTAACCCGGGCGCGTATCCGCCGAACCCGTATATGGGTCAGCCAATGTGGCAACAACCA

550 570 590
GGACCTGAGCAGCAGGATCCTGACAATTAGCTTGGCCTAATAAACTAGCCGTCTAATTCG

36

```
     610                        630                        650
AAGCTCTCCCCGGTGGATCTACTCAAGAAGAATGTTAATAGAAAACTATTGCGACAT

     670                        690                        710
AAAAAGTTTGGTGTAGTAGAATAATTTCTGTTTTATGATCCATGGATTTATCTATTGTTA

     730                        750                        770
TTCAGTTTGGTTTATCTTGTCATCAAACTGTTTTCGGTCAATGTAACAAATTCATAAACT

     790                        810                        830
GAGAATTGAACTTACAAAAGGCTAGATTACTACTTATAAAGTTCAAAGCTAAAAAAAAAA

AAAAAAAA
```

SEQ ID NO: 89

Amino acid sequence of T

```
ArgAspHisProAlaTyrHisGlnIleHisGlnGlnGlnGlnGlnLeuThrGlnGln
                              10                           20

LeuGlnSerPheTrpGluThrGlnPheLysGluIleGluLysThrThrAspPheLysAsn
                              30                           40

HisSerLeuProLeuAlaArgIleLysLysIleMetLysAlaAspGluAspValArgMet
                              50                           60

IleSerAlaGluAlaProValValPheAlaArgAlaCysGluMetPheIleLeuGluLeu
                              70                           80

ThrLeuArgSerTrpAsnHisThrGluGluAsnLysArgArgThrLeuGlnLysAsnAsp
                              90                          100

IleAlaAlaAlaValThrArgThrAspIlePheAspPheLeuValAspIleValProArg
                              110                         120

GluAspLeuArgAspGluValLeuGlyGlyValGlyAlaGluAlaAlaThrAlaAlaGly
                              130                         140

TyrProTyrGlyTyrLeuProProGlyThrAlaProIleGlyAsnProGlyMetValMet
                              150                         160

GlyAsnProGlyAlaTyrProProAsnProTyrMetGlyGlnProMetTrpGlnGlnPro
                              170                         180

GlyProGluGlnGlnAspProAspAsn
```

37

SEQ ID NO: 90

Nucleic acid sequence of X

10 30 50
AGATTCGCTATTCCTGGCAAAGAAAGACAAGATTCTGTTTACAGTGGACTTCAGGAAATC

70 90 110
GATGTGAACTCTGAGCTTGTTTGTATCCACGACTCTGCCCGACCATTGGTGAATACTGAA

130 150 170
GATGTCGAGAAGGTCCTTAAAGATGGTTCCGCGGTTGGAGCAGCTGTACTTGGTGTTCCT

190 210 230
GCTAAAGCTACAATCAAAGAGGTCAATTCTGATTCGCTTGTGGTGAAAACTCTCGACAGA

250 270 290
AAAACCCTATGGGAAATGCAGACACCACAGGTGATCAAACCAGAGCTATTGAAAAAGGGT

310 330 350
TTCGAGCTTGTAAAAAGTGAAGGTCTAGAGGTAACAGATGACGTTTCGATTGTTGAATAC

370 390 410
CTCAAGCATCCAGTTTATGTCTCTCAAGGATCTTATACAAACATCAAGGTTACAACACCT

430 450 470
GATGATTTACTGCTTGCTGAGAGAATCTTGAGCGAGGACTCATGAGATATTATATCATTT

490 510 530
ACTTAGTAAGAAGACGTGTCAAGGGTATGCATGAAAAATGTTTTATTGAAATCTTTGCAT

550 570 590
CCTAGTTTGGTGGTTTATAAAATGTGCAAGATAATTGTTTCACTGAAAACTACTTGCTGT

610 630 650
GAATATGGATTCGAACAGAGCCAATTCGAAGTAGAATTTGCATATTGTAAAAAAAAAAAA

670
AAAAAAAAAA

SEQ ID NO: 91

Amino acid sequence of X

ArgPheAlaIleProGlyLysGluArgGlnAspSerValTyrSerGlyLeuGlnGluIle
10 20

AspValAsnSerGluLeuValCysIleHisAspSerAlaArgProLeuValAsnThrGlu
30 40

AspValGluLysValLeuLysAspGlySerAlaValGlyAlaAlaValLeuGlyValPro
50 60

38

```
AlaLysAlaThrIleLysGluValAsnSerAspSerLeuValValLysThrLeuAspArg
                        70                                 80

LysThrLeuTrpGluMetGlnThrProGlnValIleLysProGluLeuLeuLysLysGly
                        90                                100

PheGluLeuValLysSerGluGlyLeuGluValThrAspAspValSerIleValGluTyr
                       110                                120

LeuLysHisProValTyrValSerGlnGlySerTyrThrAsnIleLysValThrThrPro
                       130                                140

AspAspLeuLeuLeuAlaGluArgIleLeuSerGluAspSer
                       150
```

THE CLAIMS

What is Claimed Is:

1. An isolated DNA sequence that encodes a LSD1 polypeptide.

2. The isolated DNA sequence of claim 1, wherein the sequence is selected from the group consisting of SEQ ID NO13, SEQ ID NO 14 and SEQ ID NO 15.

3. The isolated DNA sequence of claim 1, wherein the sequence comprises SEQ ID NO 13.

4. The isolated DNA sequence of claim 1, wherein the sequence comprises SEQ ID NO 14.

5. The isolated DNA sequence of claim 1, wherein the sequence comprises SEQ ID NO 15.

6. The isolated DNA sequence of claim 1, wherein the DNA is cDNA.

7. The isolated DNA sequence of claim 1, wherein the DNA is genomic.

8. The isolated DNA sequence of claim 1, wherein the polypeptide comprises SEQ ID NO 16.

9. The isolated DNA sequence of claim 1, wherein the polypeptide comprises SEQ ID NO 17.

10. A protein encoded by the isolated DNA sequence of claim1.

11. A chimeric construction comprising a promoter sequence and a DNA sequence according to claim 1.

12. A transformation vector comprising the isolated DNA sequence of claim 1.

13. A mutated DNA sequence derived from the DNA sequence of claim 1.

14. A transgenic plant expressing *LSD1* mutant genes that affect resistance to herbicidal compounds that normally result in plant cell death.

15. A transgenic plant expressing *LSD1* mutant genes which affect resistance to plant pathogens that normally result in plant cell death.

16. A messenger RNA encoding LSD1.

17. An isolated DNA sequence that encodes the zinc finger consensus selected from the group consisting of SEQ ID NOS 1-3.

18. A protein containing a zinc finger protein selected from the group consisting of CxxCxRxxLMYxxGASxVxCxxC, CxxCRxxLMYxxGASxRxVxCxxC, CxxCxxLLMYxxGAxSxCxxC, CxxCxxLLxYxxGxxxVxCSSC, CSGCRNLLMYPVGATSVCCAVC, CGGCHTLIMYIRGATSVQCSCC, CGNCMMLLMYQYGARSVKCAVC, CGSCRRLLSYLRGSKHVKCSSC, and CNNCKLLLMYPYGAPAVRCSSC, wherein x is any substituted amino acid.

19. A gene encoding a zinc finger protein according to claim 18.

20. An isolated DNA sequence encoding a protein according to claim 18.

21. A recombinant plant transformed with the DNA sequence as claimed in claim 1.

22. A recombinant plant transformed with the DNA sequence as claimed in claim 20.

23. An isolated DNA molecule that hybridizes under hybridization conditions to a DNA sequence as claimed in claim 1.

24. An isolated DNA molecule that hybridizes under hybridization conditions to a DNA sequence as claimed in claim 20.

25. An isolated DNA sequence that encodes a LSD1 homologue.

26. The isolated DNA sequence of claim 25, wherein the homologue is selected from the group consisting of LOL1 and LOL2.

27. The isolated DNA sequence of claim 25, wherein the homologue is selected from the group consisting of SEQ ID NO:48, SEQ ID NO:55, SEQ ID NO:60 and SEQ ID NO:62.

28. The isolated DNA sequence of claim 25, wherein the sequence is selected from the group consisting of SEQ ID NO:47, SEQ ID NO:54, and SEQ ID NO:59.

29. The isolated DNA sequence of claim 25, wherein the sequence comprises SEQ ID NO 47.

30. The isolated DNA sequence of claim 25, wherein the sequence comprises SEQ ID NO 54.

31. The isolated DNA sequence of claim 25, wherein the sequence comprises SEQ ID NO 59.

32. The isolated DNA sequence of claim 25, wherein the DNA is cDNA.

33. The isolated DNA sequence of claim 25, wherein the DNA is genomic.

34. A recombinant plant transformed with the DNA sequence as claimed in claim 25.

35. An isolated DNA molecule that hybridizes under hybridization conditions to a DNA sequence as claimed in claim 25.

36. A protein encoded by the isolated DNA sequence of claim 25.

37. A chimeric construction comprising a promoter sequence and a DNA sequence according to claim 25.

38. A transformation vector comprising the isolated DNA sequence of claim 25.

39. A mutated DNA sequence derived from the DNA sequence of claim 25.

40. A transgenic plant expressing *LOL1* mutant genes that affect resistance to herbicidal compounds that normally result in plant cell death.

41. A transgenic plant expressing *LOL1* mutant genes which affect resistance to plant pathogens that normally result in plant cell death.

42. A messenger RNA encoding LOL1.

43. A transgenic plant expressing *LOL2* mutant genes that affect resistance to herbicidal compounds that normally result in plant cell death.

44. A transgenic plant expressing *LOL2* mutant genes which affect resistance to plant pathogens that normally result in plant cell death.

45. A messenger RNA encoding LOL2.

46. A nucleic acid that interacts with LSD1, selected from the group consisting of the nucleic acid sequences set forth in SEQ ID NOS:66-91.

47. A protein encoded by a nucleic acid according to claim 46.

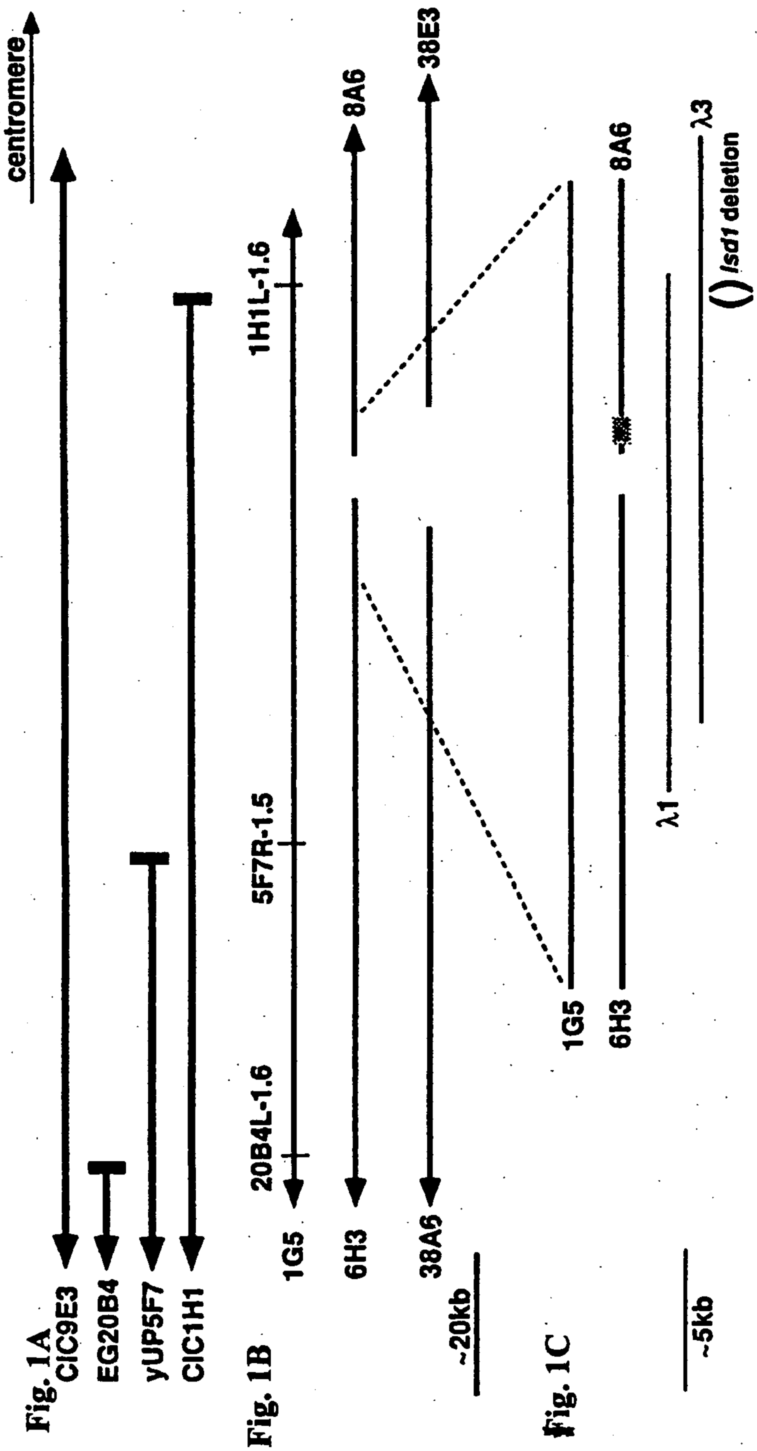

Fig. 1A

Fig. 1B

Fig. 1C

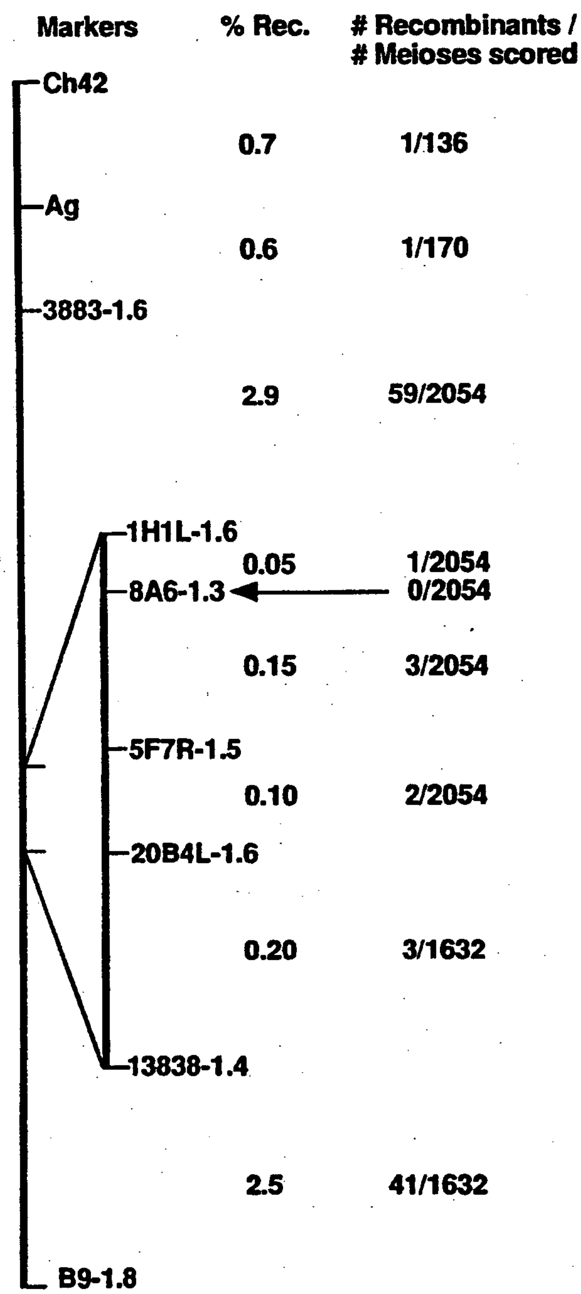
Markers
% Rec.
# Recombinants /
# Meioses scored
Ch42
0.7
1/136
Ag
0.6
1/170
3883-1.6
2.9
59/2054
1H1L-1.6
0.05
1/2054
8A6-1.3
0/2054
0.15
3/2054
5F7R-1.5
0.10
2/2054
20B4L-1.6
0.20
3/1632
13838-1.4
2.5
41/1632
B9-1.8

Figure 2

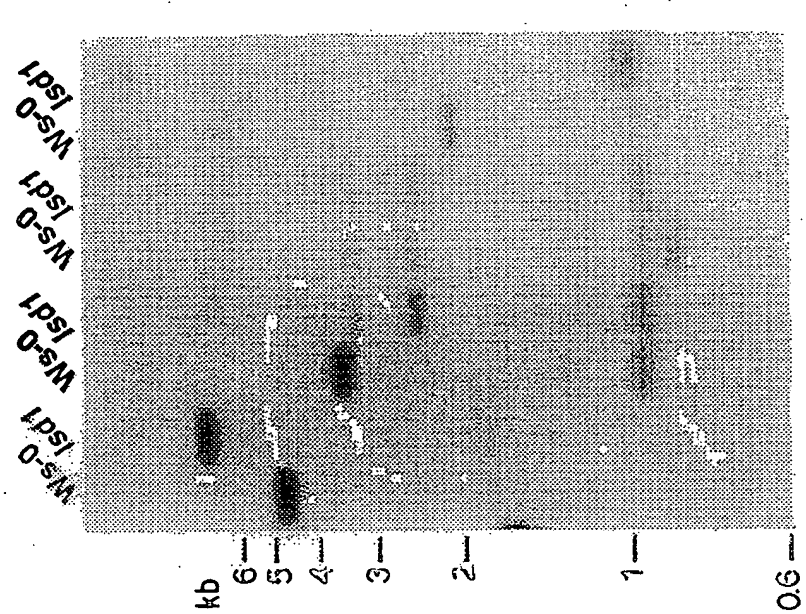
Ws-0
lsd1
Ws-0
lsd1
Ws-0
lsd1
Ws-0
lsd1
kb
6
5
4
3
2
1
0.6

FIG. 3B

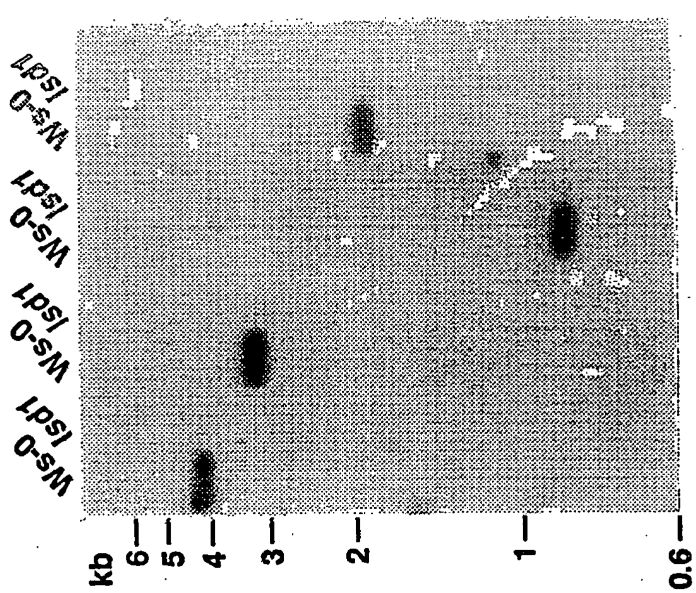
Ws-0
lsd1
Ws-0
lsd1
Ws-0
lsd1
Ws-0
lsd1
kb
6
5
4
3
2
1
0.6

FIG. 3A

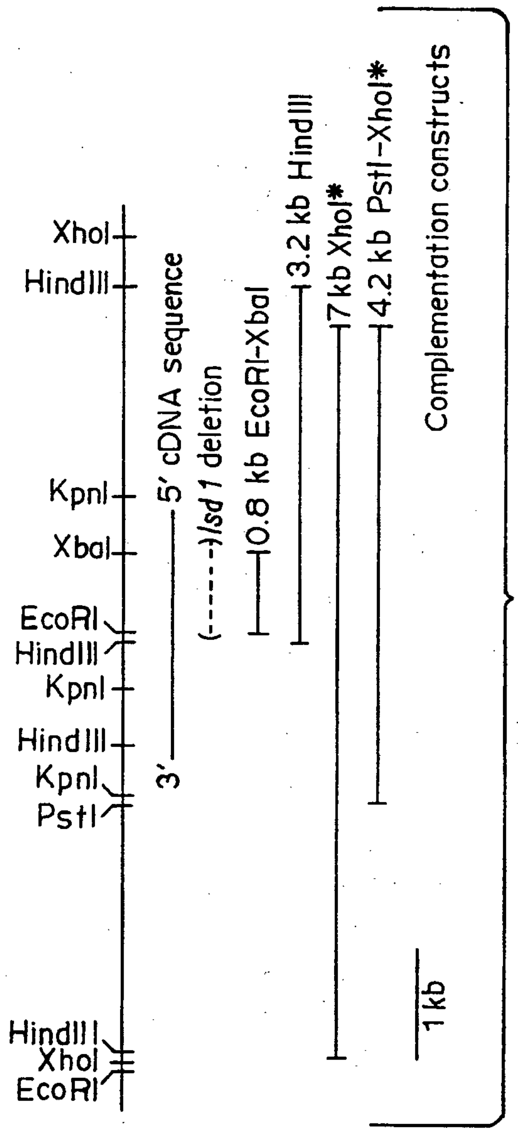
XhoI
HindIII
KpnI
XbaI
EcoRI
HindIII
KpnI
HindIII
KpnI
PstI
HindIII
XhoI
EcoRI
5' cDNA sequence 3'
(-----) lsd 1 deletion
0.8 kb EcoRI-XbaI
3.2 kb HindIII
7 kb XhoI*
4.2 kb PstI-XhoI*
1 kb
Complementation constructs

FIG. 3C

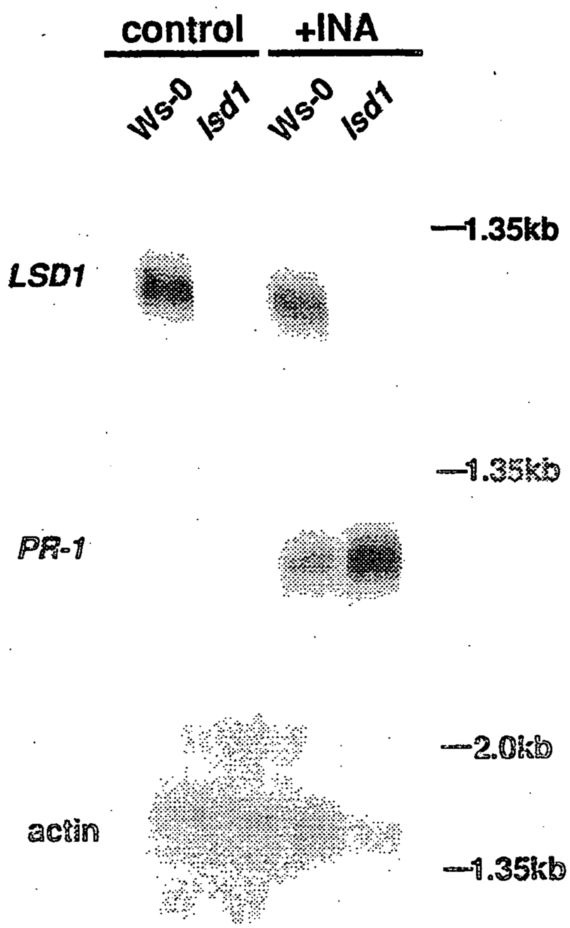
control
+INA
Ws-0
lsd1
Ws-0
lsd1
LSD1
—1.35kb
PR-1
—1.35kb
actin
—2.0kb
—1.35kb

FIG. 4

```
            10    LVCHGCRNLLMYPRGASNVRCALCNTINMV    39
                  ::| ||| :||| |||| |||  | | |:|
            51    IICGGCRTMLMYTRGASSVRCSCCQTTNLV    80
                  | || ||| |||  |||||:|  ||    |
            98    INCGHCRTTLMYPYGASSVKCAVCQFVTNV   127

Consensus:        L C  CR  LMY  GAS V C  C     V
                  I                  +    @
```

Figure 5

Fig. 6A

```
LSD1                 129 MSNGRV-PLPTNRP-NGTACPPST-STSTPPSQTQTVVVENPMSVDESGKLVSNV 180
                         || | | |:| || ||  | |  | | | |     :  :|  ||  | ::
IRS-1                611 MSPG-VAPVPSNRKGNGDYMPMSPKSVSAP-QQIINPIRRHPQRVDPNGYMM 660  (X69722)
insulin receptor
substrate
```

Fig. 6B

```
LSD1                 134 VPLPTNRP-NGTACPPSTSTSTPPSQTQTVVVENPMSVDESGKLVSNV 180
                         |||| | |   |   ||| | ||     ||
Hu Zn-finger         266 VPLPANNPVV-TTVVPSTPPSQPPAVCPPVV 295                   (S81439)
EGRα

LSD1                 134 VPLPTNRP-NGTACPPSTSTSTPPSQTQTVVVENPMSVDESGKLVSNV 180
                         |||| | |   |   ||| | ||     ||
Hu Zn-finger         277 VPLPANNPVV-TTVVPSTPPSQPPAVCPPVV 306                   (U21847)
TGF-β early induced

LSD1                 134 VPLPTNRP-NGTACPPSTSTSTPPSQTQTVVVENPMSVDESGKLVSNV 180
                         :|: ||   ||| ||| | |  || |
X.l. h-l-h           233 IPVYTNSNV-GTALPPSVSPSVSPSVT 258                       (U36194)

LSD1                 134 VPLPTNR-PNGTACPPSTSTSTPPSQTQTVVVENPMSVDESGKLVSNV 180
                         | ||| |  | | | ||| | || ||
Hu ELK1              179 VVLP-NAAPAGAAAPPSGSRSTSPS 202                         (M25269)
```

Fig. 6C

```
LSD1                 130 SNGRVPLPTNRPN-GTACPPSTSTSTPPSQTQTVVVENPMSVDESGKLVSNV 180
                         | : ||:|  || | | |       | | |  :   |:|| |:: ||  |
Z.m. IBP             548 SRALVPVPAADPNAG-AIVPANKSKRSPEQGQRRIRR-PFSVAEVEALVLAV 597  (S46309)
```

Fig. 7A

```
LSD1        10    LVCHGCRNLLMYPRGASNVRCALCNTINMV
            51    IICGGCRTMLMYTRGASSVRCSCCQTTNLV
            98    INCGHCRTTLMYPYGASSVKCAVCQFVTNV

consensus          C  CR  LMY  GAS V C  C     V
```

Fig. 7B

```
LOL1        35    LVCSGCRNLLMYPVGATSVCCAVCNAVTAV
            74    LVCGGCHTLLMYIRGATSVQCSCCHTVNLA
           112    VNCGNCMMLLMYQYGARSVKCAVCNFVTSV

consensus          C  C  LLMY  GA SV C  C  V
```

Fig. 7C

```
LOL2        61    MVCGSCRRLLSYLRGSKHVKCSSCQTVNLV
            99    VNCNNCKLLLMYPYGAPAVRCSSCNSVTDI

consensus          C  C LL Y  G   V CSSC  V
```

Fig. 8A

```
First zinc finger
LSD1          LVCHGCRNLLMYPRGASNVRCALCNTINMV
LOL1          LVCSGCRNLLMYPVGATSVCCAVCNAVTAV
LOL2          MVCGSCRRLLSYLRGSKHVKCSSCQTVNLV

consensus      VC  CR LL Y  G   V C  C     V
```

Fig. 8B

```
Second zinc finger
LSD1          IICGGCRTMLMYTRGASSVRCSCCQTTNLV
LOL1          LVCGGCHTLLMYIRGATSVQCSCCHTVNLA
LOL2          VNCNNCKLLLMYPYGAPAVRCSSCNSVTDI

consensus         C  C   LMY  GA  V CS C
```

Fig. 8C

```
Third zinc finger
LSD1          INCGHCRTTLMYPYGASSVKCAVCQFVTNV
LOL1          VNCGNCMMLLMYQYGARSVKCAVCNFVTSV

consensus      NCG C   LMY YGA SVKCAVC FVT V
```

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/04077

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :435/320.1; 530/300, 324, 325, 326, 350; 536/23.6; 800/205
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1; 530/300, 324, 325, 326, 350; 536/23.6; 800/205

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	DIETRICH et al. A Novel Zinc Finger Protein Is Encoded by the Arabidopsis LSD1 Gene and Functions as a Negative Regulator of Plant Cell Death. Cell. 07 March 1997, Vol. 88, pages 685-694, see entire document.	1-24
A	YANAGISAWA S. A novel DNA-binding domain that may form a single zinc finger motif. Nucleic Acids Research. 11 September 1995, Vol. 23, No. 17, pages 3403-3410, see entire document.	1-47
A	PUTTERILL et al. The CONSTANS Gene of Arabidopsis Promotes Flowering and Encodes a Protein Showing Similarities to Zinc Finger Transcription Factors. Cell. 24 March 1995, Vol. 80, pages 847-857, see entire document.	1-47

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 MAY 1998	23 JUN 1998

Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer TERRY A. MCKELVEY Telephone No. (703) 308-0196

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LIPPUNER et al. Two Classes of Plant cDNA Clones Differentially Complement Yeast Clacineurin Mutants and Increase Salt Tolerance of Wild-type Yeast. The Journal of Biological Chemistry. 31 May 1996, Vol. 271, No. 22, pages 12859-12866, see entire document.	1-47

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/04077

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

A01H 5/00, 7/00, 9/00, 11/00; C07K 7/08, 14/415; C12N 15/29, 15/63

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, DIALOG
search terms: LSD1, LOL1, LOL2, zinc finger, lesions simulating disease resistance, transcription factor, plant, cell death, LSD one like,

THIS PAGE BLANK (USPTO)